Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Stanford University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
JAMES DEESE, *Johns Hopkins University*
PAUL M. FITTS, *University of Michigan*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
HAROLD W. HAKE, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
HERSCHEL W. LEIBOWITZ, *International Business Machines Corporation, Bethesda, Maryland*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
LEO POSTMAN, *University of California, Berkeley*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*
DELOS D. WICKENS, *Ohio State University*

CONTENTS

Use of Temperature Stress with Cool Air Reinforcement for Human Operant Conditioning: G. L. PAUL, C. W. ERIKSEN, AND L. G. HUMPHREYS 329
Simultaneous Contrast as a Function of Test-Field Area: A. L. DIAMOND 336
Mediated Satiation in Verbal Transfer: L. A. JAKOBOVITS AND W. E. LAMBERT 346
Work Decrement and Reminiscence in Pigeon Operant Responding: C. A. BONEAU AND S. AXELROD 352
Studies of Distributed Practice: XXII. Some Conditions Which Enhance Retention: B. J. UNDERWOOD, G. KEPPEL, AND R. W. SCHULZ 355
The Role of Response Similarity in Proactive Inhibition: K. M. DALLETT 364
Effects of Nonreinforced Trials in Two-Choice Learning with Noncontingent Reinforcement: J. G. GREENO 373
Retention of First-List Associations as a Function of the Conditions of Transfer: L. POSTMAN 380
Effects of Nonreinforcement on Subsequent Reinforced Running Behavior: S. J. HARRIS, M. G. SMITH, AND S. WEINSTOCK 388
Effects of Orienting Task, Practice, and Incentive on Simultaneous Incidental and Intentional Learning: A. MECHANIC 393
Retroactive Inhibition of R-S Associations: G. KEPPEL AND B. J. UNDERWOOD 400
Cue Selection in Paired-Associate Learning: B. J. UNDERWOOD, M. HAM, AND B. EKSTRAND 405
Consummatory and Instrumental Responding as Functions of Deprivation: G. COLLIER 410
Depth Perception in Rotating Dot Patterns: Effects of Numerosity and Perspective: M. L. BRAUNSTEIN 415
Supplementary Report: Proactive Inhibition as a Function of the Method of Reproduction: J. L. WIPF AND W. B. WEBB 421
Supplementary Report: Effect upon Sensory Preconditioning of Backward, Forward, and Trace Preconditioning Training: J. D. WYNNE AND W. J. BROGDEN 422
Supplementary Report: Semantic Generalization in Probability Learning: J. P. DAS 423

American Psychological Association

Vol. 64, No. 4 October 1962

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $10.00, or $20.00 annually. Single copies are $2.00. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, **53**, 1–2).

Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association*. When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1962 the publication cost is $30 per printed page.

All **regular** articles, **early publication** articles, **supplementary reports,** and **replication reports** should be submitted to the Editor-elect: David A. Grant: Department of Psychology, University of Wisconsin; Madison 6, Wisconsin. Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts may be obtained from the Editor or from the APA Central Office.

HELEN ORR, Managing Editor | ELIZABETH S. REED, Advertising Manager | RUTH P. JAGOLINZER, Editorial Assistant

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1962 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 64, NO. 4 OCTOBER 1962

USE OF TEMPERATURE STRESS WITH COOL AIR REINFORCEMENT FOR HUMAN OPERANT CONDITIONING[1]

GORDON L. PAUL, CHARLES W. ERIKSEN, AND LLOYD G. HUMPHREYS

University of Illinois

The attempts that have been made to investigate operant conditioning in normal human *S*s have dealt primarily with verbal behavior. The general approach has been to manipulate the frequency of usage of word classes using as a reinforcement various signs of social approval such as "umhm," "good," etc. However, Buchwald (1960) has raised doubts as to the effectiveness and unambiguity of these signs of social approval as reinforcers and other *E*s (Eriksen, 1960; Krieckhaus & Eriksen, 1960; Levin, 1961) have found that changes in *S*'s verbal behavior do not convincingly occur in the absence of *S*'s ability to verbalize relevant intentions and hypotheses concerning relationships between the experimental variables and changes in his behavior.

Attempts to obtain operant conditioning of nonverbal responses have been reported by Verplanck (1955, 1956) and Hefferline, Keenan, and Harford (1959). However, Verplanck's results have come under serious question as a result of an attempt to replicate his procedures reported by Azrin, Holz, Ulrich, and Goldiamond (1961) and the results of Hefferline et al. are difficult to interpret owing to inadequate questioning of *S*s for relevant verbalizable hypotheses.

Interpretation of these studies is also complicated by the failure of the investigators to include appropriate control groups. To show clearly that operant conditioning occurred, these studies would have required a control group, treated identically with the experimentals and receiving the same number of reinforcements with the exception that the reinforcements for the controls would have been randomly administered and not contingent upon any particular response.

In the study reported below we have attempted to remedy the deficiencies in the previous attempts at operant conditioning of nonverbal responses in human *S*s. By the use of a new technique we were able to administer positive primary reinforcement to human *S*s under conditions directly comparable to those obtaining in animal experimentation. Three

[1] This research was supported by Mental Health Grant M-1206, National Institutes of Health, Public Health Service.

different classes of motor responses were studied and particular care was taken to assess the extent to which behavior modification depended upon awareness as signified by *S*'s ability to verbalize relevant mediational steps that intervened between the experimental variables and his behavior.

METHOD

Subjects.—Fifty-two females ranging in age from 18 to 30 yr. were obtained on a volunteer basis from elementary classes in psychology, education, and speech during the summer session at the University of Illinois.[2] As an added inducement, 5 dollars in services at a local beauty salon was offered volunteers.

Procedure.—Subjects were run individually. After *S* had changed from street clothes to a playsuit or swimsuit she was told that this was an experiment on the effects of heat and isolation on psychological processes. Further, that the study was government supported and was being conducted as part of a "lady in space project." The *S*s were then shown a commercial speed temperature graph demonstrating the extremely high ram air temperature (1500°) recorded on the surface of rockets. The graph also stated that "experiments by the University of California at Los Angeles have shown that man can withstand a temperature of 250°." They were then told that since this study was only concerned with psychological reactions, an increase in relative humidity had been substituted for an increase in temperature above a certain point in order to remove any physical risk. During this preliminary period *E* also determined whether *S* had any physical restriction which would be a basis for exclusion from the study and then gave *S* a brief synopsis of what she would do for the rest of the period.

Next, *S* was instructed in the operation of a stationary ergocycle which she rode for 5 min. She was told the purpose of this exercise was to bring her body temperature up so that it would not be such a drastic change when she entered the "space" chamber. After riding the ergocycle *S* was immediately directed into the chamber (maintained at a temperature of 105° F. and relative humidity of 85%). She was told to sit with her face directly in front of an insulated pipe which protruded into the chamber since cool air, high in oxygen content, would be blown through the pipe from time to time.

The chamber door was then closed and bolted. Further conversation was by means of an intercommunication system with *S* being observed through a one-way mirror in one porthole of the chamber. The *S* was told that she should ask any questions she might have immediately following the instructions, since no communication could take place once the "sonar" signals had begun. A standard set of instructions was then read to *S* further explaining the "increasing" heat in the chamber. The *S* was told: "You will probably find the air most beneficial if you breathe it directly." Also included were detailed instructions on the performance of a pseudo task consisting of differentiating between sonar signals and forming patterns with colored washers on a peg board: "Make as many different designs as you can. If you drop a washer, replace it with one of appropriate color from the stack on your right. Completely fill one-half of the board and then work back to the other half. Continue this process until I tell you the time is up. Are there any questions?"

Approximately 5 min. elapsed from the time *S* first entered the chamber through completion of the instructions. Following the instructions a time clock and tape recorder which presented the signals for the pseudo task were started. The *E* observed *S* for the following 5 min. and determined the specific response within the various response classes which was to be conditioned.

The recording apparatus was then started and the initial rate of the operant was recorded by throwing a toggle switch each time the selected response occurred. After 5 min. of recording the operant level, a switch was thrown which connected a programmer and timer into the circuit marking the beginning of the conditioning phase. During this phase which lasted for 35 min., 10 sec. of cool air (primary reinforcement) and a red light (secondary reinforcement) were automatically timed and administered on a programed schedule activated by the same switch which recorded responses.

A decreasing ratio reinforcement schedule was used in which the first 20 responses received 1:1 reinforcement, the next 15 responses received 2:3 reinforcement, the next 14 received 1:2 reinforcement, and the remaining responses received 1:3. The latter schedule applies to responses in which only

[2] Female *S*s were used since several males (run on a pilot basis), to demonstrate the cultural stereotype of "masculinity," refused air although they greatly desired it. This confounding was rare in females.

one program step could be made in every 10-sec. period. The record of responses on the other hand was kept on each specific response regardless of time interval with the exception of single responses lasting for a period of more than 10 sec. which were then recorded as two responses. This procedure was adapted to maximize unawareness.

Following the 35 min. of conditioning, the programmer and timer were disengaged and a 10-min. extinction phase was initiated in which responses were recorded as above.

Upon completion of the conditioning phase, *S* was assisted out of the chamber and escorted directly to a shower room, with instructions to return to the laboratory for a short interview.

Control *Ss* were tested for each response class, receiving exactly the same treatment as experimental *Ss* with the exception of the conditioning phase. During this 35-min. period a technician operated a separate switch which activated the programmer and timer that delivered rewards. These rewards were administered in accordance with a master program which had been prepared from the average number of program reinforcements over time that had been delivered to the experimental *Ss* in the same response class. Thus the control *Ss* received essentially the same number of rewards at the same times as the experimental *Ss* but the rewards were not contingent upon any specific response of theirs. The record of responses was taken in the same manner for both experimental and control *Ss*.

Upon *S*'s return from the shower room the following questions were asked and answers recorded verbatim:

1. Would you tell me, in your own words, exactly what we're studying?
2. Did you like or look forward to the cool air and oxygen? Relatively how much?
3. Were you able to concentrate on the task the entire period?
4. How well do you believe you did?
5. Do you feel you had control of the cool air at any time? How? When?
6. Would you mind telling me specifically what you thought about when you were in the chamber.
7. Did anything you did at any time during the period have any influence upon when the light or air came on? When?
8. If I were to tell you that something you did determined when the air came on what would you guess it might be?
9. Would you describe any emotions you may have experienced during the period.

Running notes of *S*'s behavior in the chamber were kept, specifically noting actions correlated with the operant or onset of reward. A record was also kept of *S*'s performance on the pseudo task to provide some "feedback" and in combination with the buffer items in the postexperimental interview to maintain face validity of the "space" experiment.

The reliability of *E* in recording responses was determined for each of the three response classes by having 2 other *Os*, the environmental laboratory technician and a graduate student in psychology, simultaneously observe *Ss* for an entire period and record the operant responses with paper and pencil by 1-min. intervals.

Three general categories of responses were studied: hand movements (Group H), face and mouth movements (Group FM), and foot movements (Group F). Subjects were randomly assigned to one of the response class groups with the specific response in this category that was to be reinforced (e.g., Group H, touch chin with right hand; Group FM, press lips; Group F, tap ball of right foot on deck) determined during the 5-min. observation period that preceded the determination of the operant level of the response. Twelve *Ss* were run in each response category with an additional 4 control *Ss* for each category. In the course of experimentation an additional 4 *Ss* were excluded: 2 because the experimental period had to be terminated before completion; 1 because she did not like the reinforcement; and 1 who immediately became aware of the contingency of the reinforcement and her response and stopped responding because she "thought it would ruin the experiment." No *Ss* were initially aware of the true nature of the study.

Apparatus.—The experiment was conducted in the Physical Environment Laboratory at the University of Illinois. The laboratory housing all equipment was maintained at a temperature of 70° F. and 35% relative humidity.[3]

The ergocycle which *Ss* rode prior to entrance into the chamber was the Illinois electrodynamic bicycle ergometer with an armature load of 2400 ft-lb at 50 rpm. The chamber was a 6-man low pressure chamber soundproof and cork insulated with cylindrical

[3] Lawrence Siler, technician in charge of the Environmental Unit, maintained a vigilant watch on all equipment used, assuring constant environmental conditions throughout the experiment.

TABLE 1

RELIABILITY COEFFICIENTS BETWEEN *E* AND 2 *O*s BASED ON SUCCESSIVE 5-MIN. PERIODS

Group	EO_1	EO_2	O_1O_2
H	.999	.998	.999
FM	.947	.915	.947
F	.972	.923	.965

external dimensions of 12 × 7 ft., built for the United States Air Force by the Pittsburgh Des Moines Steel Company. The internal compartment was a 6-ft. sphere. The environmental controls were of the Johnson Service Company. Conditions inside the chamber were maintained at a temperature of 105° F. (+ or −2°) and relative humidity of 85% (+ or −5%). The air velocity within the chamber determined by an aneotherm air meter varied from 10 to 18 ft. per min. at different points.

Rewards were administered through a 1-in. insulated pipe, 4 ft. long. Twin blowers drew in air from the laboratory (70° F. and 35% humidity) and blew it through the pipe, emerging at *S*'s face at 78° F. with a velocity of 710 ft. per min. Operant responses were recorded on a Hunter stylus recorder which was wired directly to the control switch. The mechanism which timed and administered rewards, also wired in series to the control switch, included a Hunter Model III electronic timer and a Ridgely automatic programmer. A 25-w. red light inside the chamber was wired to operate simultaneously with the activation of the blowers and to serve as a secondary reinforcer filling the short time interval between activation of the blowers and arrival of the cool air on *S* (less than 1 sec.).

The apparatus for the pseudo task consisted of a 1 × 3 ft. pressed wood board half of which was painted white with 42 bolts mounted on its face. The bolts contained seven each of three different colored washers. The sonar signals consisted of a magnetic tape of variable signals ranging between 800 and 2000 kc. as recorded from a variable oscillator. The tape was presented over a tape recorder which was wired into the intercommunication system with which the chamber was equipped.

RESULTS

Before turning to the results the reliability of *E* in recording the occurrence of responses needs to be documented. Table 1 presents the product-moment coefficients obtained between 2 independent *O*s and *E*. These figures were based on a single *S* from each experimental group with the total responses in successive 5-min. periods as the basic unit.

On the basis of *S*s' verbalizations during the postexperimental interview, they were classified into aware and unaware subgroups within each response category immediately upon completion of questioning. Those *S*s were classified as aware who gave an affirmative answer to Questions 5, 7, or 8 in the postexperimental interview and who could further state the correct contingency between the operant response and the reward, or name a response the occurrence of which would have led to an increase in the operant. In the case of these latter correlated hypotheses the accuracy of *S*'s verbalization was checked against the running notes that had been made of her behavior during the experiment. Subjects who gave negative answers to these three questions or who could not state the correct or correlated contingency were classified as unaware. Table 2 shows the number of aware *S*s in each of the three response groups and also the question of the three listed which elicited the verbalization of the correct or correlated contingency.

TABLE 2

NUMBER OF AWARE *S*s IN EACH GROUP DESIGNATED BY THE QUESTION ELICITING EVIDENCE OF AWARENESS

Group	Question		
	5	7	8
H	3	2	1
FM	1	2	4
F	0	1	2

Evidence of conditioning was evaluated separately within each of the three response groups by a simple analysis of covariance (aware, unaware, and control subgroups) using the average number of responses during the last 15 min. of the conditioning phase for each subgroup adjusted according to the average number of responses occurring during the operant period. The summary of these covariance analyses is given in Table 3. The adjusted and unadjusted means are presented in Table 4.

These analyses indicate that conditioning was obtained in Groups H and FM but, as seen in Table 4, apparently only in *Ss* who became aware of the contingency between their response and the reward.

Curves of the average response rate per minute over successive 5-min. periods are shown in Fig. 1 for aware, unaware, and control *Ss* within each of the response groups. These plots reveal, without adjustment for operant level, that the two groups demonstrating conditioning show a steady increase of response for the aware *Ss* during conditioning and a decline during extinction. No such change is present for the unaware *Ss* or the controls.

There was some concern that variations in the daily environment of *Ss* might have some effect on the results since the value of the reward depended for the most part upon the effect of heat and humidity within the chamber. Temperature and humidity data for the hour on which each *S* arrived for the experiment were obtained from the Illinois State Climatologist. The ranges of the mean temperature and

TABLE 3

COVARIANCE ANALYSES ON THE MEAN RESPONSES DURING THE LAST 15 MIN. OF CONDITIONING ADJUSTED TO OPERANT LEVEL

Group	Source	*df*	*MS*	*F*
H	Subgroups	2	446.9	4.06*
	Error	12	110.1	
FM	Subgroups	2	1463.7	7.93**
	Error	12	184.5	
F	Subgroups	2	43.7	.18
	Error	12	237.5	

* $P < .05$.
** $P < .01$.

TABLE 4

MEAN OPERANT LEVEL AND MEAN AND ADJUSTED MEAN RESPONSES DURING LAST 15 MIN. OF CONDITIONING

Group	Subgroup	*N*	Operant Level		Last 15 Min. of Conditioning		
			Mean	*SD*	Obtained		Adjusted
					Mean	*SD*	Mean
H	Aware	6	5.2	1.14	30.7	12.09	29.9
	Unaware	6	4.8	2.19	15.7	8.84	15.8
	Control	4	4.5	2.69	11.8	10.92	12.8
FM	Aware	7	7.4	1.99	53.1	14.50	55.1
	Unaware	5	7.8	2.71	28.0	18.86	28.2
	Control	4	8.8	3.77	31.0	20.23	26.8
F	Aware	3	14.3	7.41	48.7	17.99	44.9
	Unaware	9	8.4	5.87	44.2	14.02	45.6
	Control	4	10.2	2.48	40.2	11.09	40.0

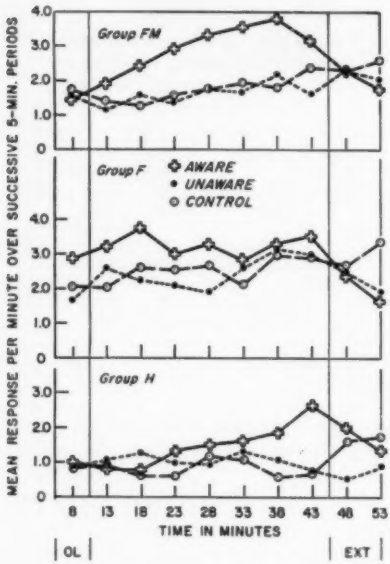

FIG. 1. Mean responses per minute over successive 5-min. periods. (The times shown on the abscissa represent midpoints of the successive 5-min. intervals from completion of instructions—see text.)

humidity for each subgroup were 79.5–83.25° F. and 43.0–52.25%, respectively. There appeared to be no differences attributable to external environmental conditions over these slight ranges.

The success of this technique and the apparatus used was vouched for by all *S*s save one. In response to Question 2, in which *S*s were asked how well they liked the cool air, responses such as: "Very, very much"; "It was ecstasy"; "Desperately"; "Like a Manhattan"; "Couldn't have survived without it" were typical. In pilot studies in which *S*s were informed as to how they could obtain the cool air, some *S*s gave as many as eight responses per min. when six responses per min. were sufficient to maintain a constant supply of cool air. It was not uncommon for *S*s to verbally request air, particularly during the extinction period, and *E* suffered some verbal abuse from *S*s in all groups when these requests were not fulfilled.

DISCUSSION

The results obtained in Groups H and FM require little interpretation. When plotted cumulatively, the extinction curves of the aware *S*s still show positive although reduced slope. This is not surprising since *S*s had been on a partial reinforcement schedule just prior to the extinction phase and the strength of the reward was great. Some confounding may also be present, since the controls in all three groups increased in response rate during extinction. Also to be noted is that operant levels did change as a part of a general change in activity during the period, exclusive of experimental manipulation, Fig. 1. The reason that *S*s in Group H were generally lower in response rate may be due to the fact that the pseudo task required those *S*s to be using their hands constantly.

The results of this study give added support to the increasing evidence against "learning without awareness." Only those *S*s who could verbalize a contingent relation between their behavior and the reinforcement demonstrated learning. Since learning did occur for these *S*s it would appear that the reinforcement was effective. Further, the fact that somewhat less than 50% of the *S*s became aware of the contingency would suggest that the task allowed room for "learning without awareness" had such a process been operative.

A finding that strikes us as quite remarkable was that over half of the *S*s did not learn. Results such as this are not surprising in studies utilizing verbal reinforcements where the "rewards" are secondary and have constituted a part of the general social background of *S*s for many years. However, in the present case the reward was primary and dramatic and gained the attention of all *S*s to such a degree that it seems surprising

that humans could spend 35 min. with the occurrence of this reward contingent upon their behavior and not learn the contingency.

A factor that may at least partially account for *Ss* not learning or becoming aware of the contingency of the reward upon their behavior is the high operant level of the responses selected for conditioning. This is particularly true in Group F where only 3 *Ss* became aware of the contingency. Since the reward lasted for 10 sec., most *Ss* in this group would receive nearly continuous reinforcement by merely maintaining their operant level. In fact Table 4 demonstrates that the mean number of responses of the control *Ss* in Group F was greater than that of the aware *Ss* in Group H.

The data forcefully demonstrate the need for appropriate control *Ss* in work on operant conditioning since it is obvious that parameters other than the experimentally controlled reinforcement exert an influence on *Ss*' behavior. Also, the thorough exploration of *Ss*' verbalized hypotheses and other behavior is shown to be an absolute requirement in determining awareness. Only 4 *Ss* of 16 who were ultimately classified as aware volunteered the correct or correlated contingency in response to the initial question.

It appears that the technique utilized in this study offers a means to replicate on humans the behavioral changes found in subhuman animals. The high temperature and humidity, supplemented by appropriate instructions, allows for the manipulation of a reinforcement which definitely seems comparable in reward value to those administered to deprived animal *Ss*. Isolation within the soundproof chamber also affords a more comparable situation, and one in which most stringent environmental and task controls are built in. Some limitations are placed on a complete survey of operant conditioning phenomena owing to the duration of the reward (determined in pilot work as the optimum), however this limitation appears to be slight when compared with the ease of administration and overall time saved by the procedure.

Summary

This study was concerned with a technique for producing operant conditioning of human motor responses with special emphasis on conditioning without awareness. Female *Ss* who thought they were participating in a "space" experiment were enclosed in a chamber at 105° F. and relative humidity of 85%. A 10-sec. stream of cool air served as a reward for operant responses. Two of the three experimental groups (designated by response class) demonstrated conditioning and intensive investigation of *Ss*' awareness revealed that conditioning was apparent only in those *Ss* who were able to verbalize mediational steps that intervened between the experimental variables and the changes in their behavior.

References

AZRIN, N. H., HOLZ, W., ULRICH, R., & GOLDIAMOND, I. The control of the content of conversation through reinforcement. *J. exp. Anal. Behav.*, 1961, **4**, 25–30.

BUCHWALD, A. M. Supplementary report: Alteration in the reinforcement value of a positive reinforcer. *J. exp. Psychol.*, 1960, **60**, 416–417.

ERIKSEN, C. W. Discrimination and learning without awareness: A methodological survey and evaluation. *Psychol. Rev.*, 1960, **67**, 279–300.

HEFFERLINE, R. F., KEENAN, B., & HARFORD, R. A. Escape and avoidance conditioning in human subjects without their observation of the response. *Science*, 1959, **130**, 1338–1339.

KRIECKHAUS, E. E., & ERIKSEN, C. W. A study of awareness and its effects on learning and generalization. *J. Pers.*, 1960, **28**, 503–517.

LEVIN, S. M. The effects of awareness on verbal conditioning. *J. exp. Psychol.*, 1961, **61**, 67–75.

VERPLANCK, W. S. The control of the content of conversation: Reinforcement of statements of opinion. *J. abnorm. soc. Psychol.*, 1955, **51**, 668–676.

VERPLANCK, W. S. The operant conditioning of human motor behavior. *Psychol. Bull.*, 1956, **53**, 70–83.

(Received September 20, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 336-345

SIMULTANEOUS CONTRAST AS A FUNCTION OF TEST-FIELD AREA[1]

A. LEONARD DIAMOND

Psychological Research Center, University of Hawaii

The phenomenon of simultaneous contrast is illustrated in the change in brightness of a visual field without a corresponding change in the luminance of that field. That is, if next to a small illuminated square, which we shall call the test field, we place another illuminated square of much greater luminance, which we shall term the inducing field, we will perceive a decrease in the brightness of the test-field square even though the test field is held constant in luminance.

A number of parameters of this phenomenon have been investigated in previous experiments. The luminance of the test and inducing fields (Diamond, 1953; Heinemann, 1955), the separation between test and inducing fields (Fry & Alpern, 1953; Leibowitz, Mote, & Thurlow, 1953), and the area of the inducing field (Diamond, 1955) have all been found to be pertinent variables in the contrast effect.

It is now of interest to know how the variation of the test-field area will affect simultaneous contrast. Specifically, a theoretical formulation (Diamond, 1960; also see Discussion below), which describes the relationships between the above-mentioned parameters, would necessarily predict little or no change in test-field brightness as the test-field area is varied in such a way that its center remains at a constant distance from the center of the inducing field. The present experiment is designed to test this prediction.

Our general method is as follows (see Fig. 1): To *S*'s right eye is presented an inducing field (*i*), a rectangle twice as wide as it is long; below the inducing is the test field (*t*), equally as wide as the inducing field but variable in height, and thus in area. To *S*'s left eye is presented the match field (*m*), which is either kept equal in size to the test field, as the test-field area is varied, or is held constant in size. (As we shall see, whether we vary the match-field size or hold it constant makes little difference in the results.) As the test-field area is varied, the distance between the centers of the test and inducing fields is held constant. The brightness of the test field is thus measured as a function of its area for three different values of test-field luminance (.69, 1.60, and 2.68 log mL.) over a wide range of inducing-field luminances ($-\infty$ to 2.79 log mL.).

METHOD

Apparatus

Description of S's view.—A modification of the apparatus used by Diamond (1955) is employed in the present experiment. The patterns seen by *S* are different for each eye (see Fig. 1).

The pattern, R, to the right eye only, includes: (*a*) a test-field rectangle (*t*), 33′ in visual angle along the horizontal, and variable in its vertical extent from zero to 33′; (*b*) above the test field with its center held at a constant distance from the center of the test field, an inducing-field rectangle (*i*), also 33′ wide but with a fixed vertical dimension of 16.5′; and (*c*) a small fixation point (P) located 21′ to the left of the test field. The

[1] This work was supported by a research grant NSF-G9588 from the National Science Foundation.

second pattern, L, presented to the left eye only, includes: (*a*) a match-field rectangle (*m*), 33′ and variable in height; and (*b*) a small fixation point (P) located 21′ to the right of the match field.

In the binocular view, *S* is instructed to fuse left- and right-eye fixation points into the point P so that the match, test, and inducing fields are held in constant position relative to one another. Stimulation may be considered essentially foveal since the visual angle between the fused fixation point and furthest corner of any field is never more than 70′.

In order to make it easier for *S* to fuse the left- and right-eye fixation points, a circular prism, the refracting angle of which is 5°, is placed in each eyepiece of the apparatus. By rotating each eyepiece, one clockwise and the other counterclockwise, *S* can optically rotate both left- and right-eye patterns such that he can vary the horizontal separation between them and thus more easily superimpose (fuse) the left- and right-eye fixation points. No systematic changes in the brightness of either right- or left-eye patterns occurred as a result of prism rotation.

Apparatus controls.—The apparatus is designed to control the following variables: (*a*) the luminances of the test, inducing, and match fields; and (*b*) the areas of the test and match fields. Luminance controls are discussed in detail by Diamond (1955). The general arrangement is as follows: The right- and left-eye patterns (L and R in Fig. 1) are presented to *S* along two separate optical paths, one to each eye. The light source of each path is a 150-w. tungsten filament projection lamp, in front of which is a section of heat-absorbing glass. The light, diffused by flashed opal glass, travels through its particular pattern (L or R). The luminance in either path can be continuously varied by fixed filters and a fixed and movable Polaroid. Made parallel by 4-diopter lenses, the light finally travels through 3-mm. artificial pupils into the eyes of *S* who is seated in a light-tight cubicle. The entire left optical path is adjustable horizontally for interpupillary distance by means of a screw arrangement attached to the optical bench.

The luminance of the match, test, and inducing fields are calibrated by means of binocular matches to fields of similar shapes and areas and whose luminances had been determined by a MacBeth illuminometer.

Test-field area is controlled by the use of six thin metal masks. In each of these masks is cut a test-field rectangle of a particular vertical dimension (5.5′, 11.0′, 16.5′, 22.0′, and 33.0′ in visual angle) and constant horizontal dimension (33.0′). An inducing-field rectangle of constant dimensions (16.5′ × 33.0′) is also cut in each mask such that its center is always at the same distance (24.75′) from the test-field center. This is illustrated as Pattern R in Fig. 1.

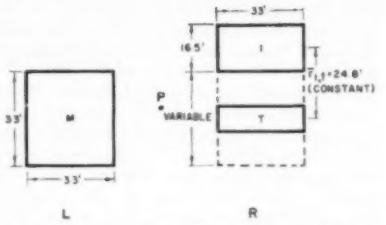

FIG. 1. The *S*'s binocular view to left (L) and right (R) eyes. (The match field—*m*—was held either at its maximum of 33 min. or equal in area to the variable test-field—*t*.)

Match-field area is controlled as follows: A thin metal mask, in which is cut a 33′ square, is placed before the diffused light source (in the left optical path). Another thin piece of straight-edged metal is mounted immediately next to the square mask such that when the straight edge of this second piece of thin metal coincides with the top edge of the match field, the match-field area is zero. When the straight-edged metal is drawn down away from the top of the match field, the match-field area increases in successive rectangular increments. In this way, the match field can be set equal in area to any of the six test-field areas.

Procedure

The experimental method was designed to investigate the brightness of the test field as a function of its area; the area was varied in six steps from 5.5′ to 33′ (vertical dimension). This entire function was studied for three test-field luminances, .69, 1.60, and 2.68 log mL., and for various inducing-field luminances as specified in Table 1.

It was desirable to determine whether variation of the match-field area would affect the results in any systematic way. Therefore, in addition to the variation of the aforementioned parameters, two conditions of match-field area were explored. That is, the match-field area was (*a*) held equal to that of the test field as the test-field area was varied and (*b*) held constant at its maximum value, i.e., 33′ square for all values of test-field area. This last condition prevailed for two inducing-

field luminances—zero luminance ($-\infty$ log mL.), and a luminance approximately equal to that of the test field, whatever it happened to be (.69, 1.60, or 2.68 log mL.).

During each experimental session, which required between 30 to 60 min., test-field area and luminance were held at constant values for that session, and both match-field area and inducing-field luminance were varied. The various match-field sizes or inducing-field luminances explored during one session were counterbalanced. Intersession periods were never shorter than 4 hr. Successive test-field areas and test-field luminances were taken in different sessions and were presented in random order.

Psychophysical method.—Seated in the light-tight cubicle, *S* initially dark adapted for 3 min., then light adapted for 3 min. to the binocular view, i.e., at a test-field luminance and a set area for a particular experimental session, the match field at the same brightness as the test field, the inducing field at a luminance set for a particular experimental point. The *S* then began making brightness-equality matches following the psychophysical method of adjustments as described by Guilford (1936). That is, *S* set the luminance of the match field so that it appeared equal to that of the test field. The *E* then changed the luminance of the match field in a random manner after each match. The *S* had to adjust the match-field luminance again until it seemed equal in brightness to that of the test field. In this manner, for each experimental point as indicated in Table 1, *S* made 10 matches. The average of these 10 match-field luminances was taken as the brightness of the test field for each experimental point. This procedure was followed for 2 *S*s, JS and RH.

Results

The data are presented for each *S* in Table 1, and averaged in Fig. 2. Figure 2 shows the log luminance of the match field (B_m) plotted as a function of the test-field area (as measured by its vertical dimension). Log B_m may be termed the brightness of the test field since its value is based upon an equality judgment between the match- and test-field brightnesses. The function is graphed for different values of test-field luminances (log B_t),

TABLE 1

Log Match-Field Luminances (mL.) for Different Test Area and Inducing and Test Luminances

Log Test Luminance (mL.)	Log Ind. Luminance (mL.)	Test Height (Min.)											
		5.5		11.0		16.5		22.0		27.5		33.0	
		RH	JS	RH	JS	RH	JS	RH	JS	RH	JS	RH	JS
0.69	$-\infty$	.73	.86	.46	.75	.57	.82	.50	.75	.51	.92	.69	.80
	$-\infty_m$[a]	.64	.99	.52	.77	.37	.88	.48	.66	.53	.81	.68	.81
	0.69	.61	.65	.64	.80	.73	.90	.48	.62	.41	.53	.61	.75
	0.70_m[a]	.60	.79	.59	.84	.41	.85	.50	.69	.50	.66	.62	.74
	1.70	.26	.50	.46	.59	.29	.34	.34	.44	.09	.36	.11	.49
	2.68	−.02	.06	.23	.30	.09	.11	−.07	.19	−.26	−.03	−.04	.22
1.60	$-\infty$	1.28	1.91	1.31	1.76	1.49	1.81	1.49	1.66	1.40	1.73	1.40	1.58
	$-\infty_m$[a]	1.75	2.01	1.47	1.79	1.28	1.70	1.37	1.54	1.51	1.85	1.50	1.65
	1.60	1.31	1.66	1.21	1.54	1.49	1.69	1.32	1.47	1.24	1.61	1.53	1.58
	1.69_m[a]	1.42	1.96	1.33	1.72	1.22	1.56	1.38	1.54	1.33	1.45	1.17	1.49
	2.23	1.33	1.70	1.25	1.46	1.18	1.38	1.18	1.37	1.39	1.53	1.26	1.75
	2.72	1.02	1.31	.99	1.17	1.01	1.30	1.18	1.14	1.02	1.11	1.26	1.45
2.68	$-\infty$	2.19	2.84	2.48	2.61	2.37	2.72	2.51	2.83	2.50	2.86	2.73	2.88
	$-\infty_m$[a]	2.59	2.99	2.80	2.94	2.46	2.76	2.45	2.81	2.51	2.85	2.65	2.92
	−0.32	2.20	2.74	2.60	2.50	2.08	2.74	2.29	2.77	2.48	2.66	2.65	2.76
	1.22	2.35	2.83	2.50	2.87	2.31	2.62	2.53	2.82	2.36	2.68	2.60	2.82
	2.68	2.30	2.75	2.47	2.62	2.38	2.47	2.49	2.72	2.27	2.61	2.66	2.70
	2.79_m[a]	2.64	2.94	2.42	2.73	2.21	2.56	2.59	2.80	2.40	2.71	2.71	2.71

[a] The subscript *m* indicates a constant maximum match-field area for all test areas.

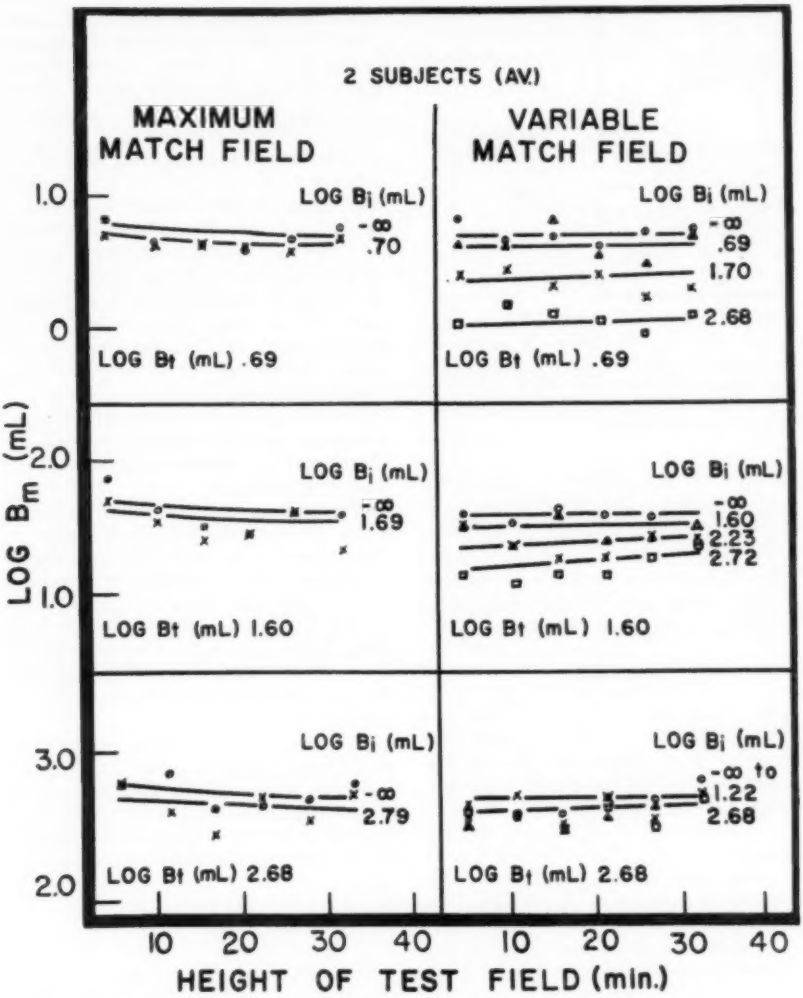

FIG. 2. Test brightness as a function of test area for different test and inducing luminances.

and for different inducing-field luminances (log B_i) and match-field areas.

The effect of test-field area upon test-field brightness is minimal. There seems to be little or no change in the brightness of the test field as its area is increased; this holds throughout the test-field and inducing-field luminance range explored and for both conditions of match-field area. Inducing luminance, however, does have an effect as seen in previous experiments (Diamond, 1953, 1955). The greater the inducing luminance,

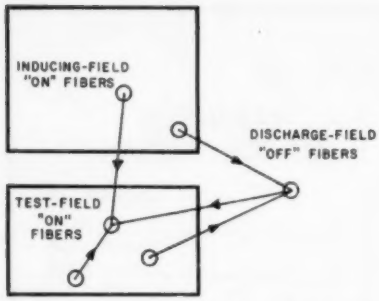

FIG. 3. Diagram of inhibitory interactions between test and inducing on fibers and spontaneously discharging off fibers. (Arrows show directions of main inhibitory effects. Circles represent any fiber or fibers in each respective field.)

greater than equality with test luminance, the lower the brightness of the test field, or the greater the depression of test brightness.

The lines drawn through the experimental points are fit by theoretical curves which can now be discussed.

DISCUSSION

The theory fit to the above data was originally devised not only to explain the phenomenon of depression seen in the present results (and its change, or lack of change, with test-area change) but also the phenomenon of enhancement (see Diamond, 1960). Under certain conditions test brightness can increase, or be enhanced, especially if surrounded by a less bright inducing or surround field. Why enhancement did not occur in the present experimental situation will be explained after the basic physiological assumptions to the theory are summarized. In Fig. 3 are diagramed the main or primary physiological events that according to our theory are basic to psychophysical depression and enhancement.

For depression of test brightness, the test field "on" fibers (which presumably mediate psychophysical brightness) are inhibited by (*a*) the inducing on fibers, (*b*) other test on fibers, and (*c*) the spontaneously discharging off fibers.

For enhancement of test brightness we must first note something not indicated in Fig. 3; i.e., off fibers exist within the test and inducing fields as well as the discharge field. Now to explain enhancement in a multiple-field situation; i.e., one in which an inducing field as well as a test field is present, we must assume that when the inducing-field luminance is zero, the off fibers immediately surrounding the test field are normally active. As the inducing field now increases from zero to some value below test-field luminance the off fiber activity within the inducing-field borders becomes diminished. Enhancement in the multiple-field situation is therefore explained as a "disinhibition" of test-field activity; i.e., inhibition by the inducing field of discharge activity releases the test field from discharge inhibition.

This, of course, is a verbal description of the theory behind the explanation of Fig. 2 data. The curves drawn through the data are mathematically determined. The mathematic function fitted to the data is based upon preliminary assumptions which we must first discuss. A more complete and detailed discussion of the assumptions is available in Diamond (1960).

Our theory then assumes a number of physiological events, concerning both on and off fibers in the retina when it is illuminated by the test circle. Most of these assumptions are based upon physiological findings in animals and are as follows.

On Fiber Frequency

The frequency f_{t_1} of a stimulated on retinal fiber (as described by Hartline, 1938):

1. directly determines the strength of the brightness response, or

$$A_t = k_1 f_{t_1} \qquad [1]$$

where A_t represents the brightness response, k_1 is a proportionality constant, and f_{t_1} is the frequency of the on fiber. This relationship has been suggested by Adrian's (1928) demonstration of the similarity between the brightness-duration curves in

man and the frequency-duration curves taken from the optic nerve of the eel.

2. is directly proportional to a power function of the luminance of the stimulating light, as suggested in Diamond's (1960) fit of Hartline and Graham's (1932) *Limulus* data, or

$$f_{t_1} = k_2 B_{t_1}^{\frac{1}{3}} \qquad [2]$$

where B_{t_2} is the luminance of light striking an on fiber.

3. is inversely proportional to the frequency of a nearby on fiber, as demonstrated by Hartline and Ratliff (1957) in the *Limulus*, or

$$f_{t_1} = \frac{k_3}{f_{t_2}} \qquad [3]$$

where f_{t_2} is the frequency of a nearby on fiber within the test field. Since according to Equation 2 above frequency is proportional to luminance then

$$f_{t_2} = \frac{k_4}{B_{t_2}^{\frac{1}{3}}} \qquad [4]$$

when B_{t_2} is the luminance striking the nearby on fiber. Within the test field this is equal to B_{t_1}.

4. is inversely proportional to the number of nearby on fibers as demonstrated by Hartline, Wagner, and Ratliff (1956) in the *Limulus*, or

$$f_{t_1} = \frac{k_5}{E_{t_2}} \qquad [5]$$

where E_{t_2} represents the total number of nearby on fibers within the test field.

5. is directly proportional to the distance between the test and nearby on fibers, as demonstrated by Hartline, Wagner, and Ratliff (1956), or

$$f_{t_1} = k_6 r_{t,t} \qquad [6]$$

where $r_{t,t}$ is the separation between the two on fibers in the test field. It becomes convenient and actually desirable, as pointed out by Diamond (1955), to combine Equations 5 and 6 such that

$$C_{t,t} = \frac{k_7 E_{t_2}}{\bar{r}_{t,t}} \qquad [7]$$

where $\bar{r}_{t,t}$ is the average separation between all test fibers and $C_{t,t}$ therefore describes the combined effect of all the individual nearby on fibers; it then follows that

$$f_{t_1} = \frac{k_8 \bar{r}_{t,t}}{E_{t_2}} = \frac{k_8}{C_{t,t}} \qquad [8]$$

Equations 2 through 8 are also applicable to the interaction between a test on fiber and inducing on fiber (see Fig. 3). Thus

$$f_{t_1} = \frac{k_9}{C_{i,t}} \qquad [9]$$

where $C_{i,t}$ represents the combined effect of all on fibers from the inducing field upon those in the test field.

6. is inversely proportional to the frequency of a nearby off fiber, as suggested in Granit's (1955) descriptions of the mutual antagonism between on and off fibers, or

$$f_{t_1} = \frac{k_{10}}{f_o} \qquad [10]$$

where f_o is the frequency of an off fiber.

7. is inversely proportional to the number of nearby off fibers at particular distances away or

$$f_{t_1} = \frac{k_{11}}{C_{d,t}} \qquad [11]$$

where $C_{d,t}$ is the combined effect of all the off fibers in the discharge field on the on fiber frequency in the test field. With respect to the frequency of an on retinal fiber in the test field, Equations 1 through 11 may be combined into the following formula:

$$f_t = \frac{K_o B_t^{\frac{1}{3}}}{K_1 B_t^{\frac{1}{3}} K_2 C_{t,t} + K_3 B_t^{\frac{1}{3}} K_4 C_{i,t} + K_5 f_o C_{d,t}} \qquad [12]$$

The subscripts of the proportionality constants (K) are chosen to coincide with those used in the more general brightness theory by Diamond (1960).

Off Fiber Frequency

The frequency, f_o, of an off fiber in the discharge field, which we assume to be spontaneously active in the dark according to data taken from the cat eye by Barlow, FitzHugh, and Kuffler (1954):

1. is directly proportional to some "internal driving force" (comparable to the external luminance effect on on fibers) which we shall assume to exist. Such a mechanism is suggested by experiments, described by Granit (1955), which show both on and off activity to increase in the retina as a result of central (reticular formation) stimulation. Thus

$$f_o = k_{12} D^{\frac{1}{3}} \qquad [13]$$

where D represents this "driving force."

2. is inversely proportional to the frequency of a nearby on test fiber, according to Granit's on-off antagonism findings cited above, or

$$f_o = \frac{k_{14}}{f_t} \qquad [14]$$

3. is inversely proportional to number and distances away of nearby on test fibers, or

$$f_o = \frac{k_{15}}{C_{t,d}} \qquad [15]$$

where $C_{t,d}$ is the combined effect of all the on fibers in the test field upon the off fibers in the discharge field. This assumption has not been tested experimentally.

4. is inversely proportional to the amount of light impinging upon the off fiber, as demonstrated by Hartline (1938). This light could be direct or scattered from the test beam stimulating the on fibers. The retinal effectiveness of scattered light in the human eye has been demonstrated by Boynton and Riggs (1951). Thus,

$$f_o = \frac{k_{16}}{B_t} \qquad [16]$$

With respect to the frequency of an off fiber in the retinal discharge field, therefore, Equations 12 through 15 may be combined into the following formula:

$$f_o = \frac{K_5 D^{\frac{1}{3}}}{K_7 B_t C_{t,d}} \qquad [17]$$

Equations 14 through 16 are also applicable to the interaction between an inducing on fiber and a discharge off fiber (see Fig. 3). Thus

$$f_o = \frac{K_5 D^{\frac{1}{3}}}{K_8 B_i C_{i,d}} \qquad [18]$$

where $C_{i,d}$ represents the combined effect of all on fibers from the inducing field upon the off fibers in the discharge field.

It should be noted that the inhibition of the discharge field by the inducing field is effectively greater when an inducing square, for example, is near the test square. This is because discharge elements near the test field are initially more effective than those far away. When the inducing field inhibits near discharge elements, the effect of spontaneous discharge on the test field is much more reduced than when the far discharge elements are inhibited. Therefore the effective discharge frequency is a function of the separation between test and inducing fields, or

$$f_o = k_{17} \frac{\bar{r}_{i,t}}{\bar{r}_{t,t}} \qquad [19]$$

where $\bar{r}_{i,t}$ is the average separation between all elements in the inducing field and those in the test field. That $\bar{r}_{i,t}$ is set in ratio to $\bar{r}_{t,t}$ is empirically required for a satisfactory fit of certain data (see Diamond, 1960, pp. 183–184). A completely rational account of the functional relationship between f_o and $\bar{r}_{i,t}$ awaits further knowledge of separate and relative effects of light scatter in the intact human eye and on-off antagonism.

Equations 17 through 19 may be combined to include the effect of the inducing field upon the off frequency, or

$$f_o = \frac{K_5 D}{K_7 B_t^{\frac{1}{3}} C_{t,d} + K_8 B_i^{\frac{1}{3}} \frac{\bar{r}_{t,t} C_{i,d}}{\bar{r}_{i,t}}} \qquad [20]$$

If we now combine Equations 1, 12, and 20 then

$$A_t = \frac{K_o B_t^{\frac{1}{3}}}{K_1 B_t^{\frac{1}{3}}(K_2 C_{t,t}) + K_3 B_i^{\frac{1}{3}} K_4 C_{i,t} + \dfrac{K_5 D^{\frac{1}{3}}(K_6 C_{d,t})}{K_7 B_t C_{t,d} + K_8 B_i \dfrac{\bar{r}_{t,t} C_{i,d}}{\bar{r}_{i,t}}}} \qquad [21]$$

The S's brightness response to the match field can be described in the same manner minus the inducing field effect, or

$$A_m = \frac{K_o B_m^{\frac{1}{4}}}{K_1 B_m^{\frac{1}{4}}(K_2 C_{m,m}) + \dfrac{K_5 D^{\frac{1}{4}}(K_6 C_{d,m})}{K_7 B_t C_{t,d}}} \qquad [22]$$

Since in the present experiment, the S is required to match the test and match field in brightness such that

$$A_t = A_m \qquad [23]$$

then if accordingly we set Equations 21 and 22 equal

$$\frac{K_o B_m^{a/4}}{K_1 B_m^{b_1/4}(K_2 C_{m,m})^{c_1} + \dfrac{K_5 D^{b_4/4}(K_6 C_{d,m})^{c_2}}{K_7 B_m C^*_{m,d}}} = \frac{K_o B_t^{a/4}}{K_1 B_t^{b_2/4}(K_2 C_{t,t})^{c_3} + K_3 B_i^{b_3/4}(K_4 C_{i,t})^{c_4} + \dfrac{K_5 D^{b_4/4}(K_6 C_{d,t})^{c_5}}{K_7 B_t C^*_{t,d} + K_8 B_i \left(\dfrac{\bar{r}_{t,t} C_{i,d}}{\bar{r}_{i,t}}\right)^f}} \qquad [24]$$

In Table 2 are presented the values of the terms in Equation 24. The procedures for determining values for the terms were as follows: The values of the luminance (B) terms were taken directly from the luminance values of the different fields. The area-separation (C) terms were calculated according to Equation 10 above. The dimensions of the discharge area were arbitrarily chosen; since only minimal inhibitory effects occur between fields separated by more than 4.5° (Fry & Alpern, 1953) of visual angle and at certainly no more than 9° (Leibowitz et al., 1953); spontaneous discharge elements further than 9° from the center of the test field were not considered to be effective.

The value of D (spontaneous discharge activity) is arbitrarily taken as the value 1. This turns out to be empirically satisfactory and must do until such time as physiological and/or psychophysical, manipulation of discharge activity reveals its actual value.

The value of the a, b, c, e, f, and K terms were determined empirically (by trial and error fitting). The value of K_o is not included in Table 2 since this term cancels out in Equation 20.

It should be noted that in Equation 20 and Table 2 are listed a total of 18 fitted constants. Were these constants chosen completely without restriction, any number of a variety of functions could be generated, so that the fit of our physiological theory to the 16 functions in our present experiment (see Fig. 2) would be meaningless. The constants, however, were not chosen arbitrarily but were those fit by Diamond (1960) to more than 45 additional functions having to do with brightness contrast and predicted by the same theory as outlined above.

As noted above, enhancement as a function of test area is not clearly evident in the data. Furthermore, although it is included as a basic component of the present theory it is not predicted for the stimulus conditions existing in the present experiment. That is, the off fibers in the discharge field (see Fig. 3) are already inhibited

TABLE 2

VALUES OF TERMS IN EQUATION 24

Determined by Experimental Situation						Empirically Fitted Values								
Luminance (Log ml.)			Area-Separation (Log)			Exponents			Proportionality Constants (Log)			"Off Drive"		
Terms	Max. Match	Variable Match	Terms	Max. Match	Variable Match	Terms	Max. Match	Variable Match	Terms	Max. Match	Variable Match	Terms	Max. Match	Variable Match
B_m	v	v	$C_{m,m}$	2.62	v	a	12.00	12.00	K_1	0.00	0.00	D	1.00	1.00
B_t	v	v	$C_{d,m}$	4.19	4.19	b_1	4.00	4.00	K_2	−2.80	−2.80			
B_i	v	v	$C_{m,d}$	1.28	v	b_2	4.00	4.00	K_3	−0.00	−0.00			
			$C_{t,t}$	v	v	b_3	4.00	4.00	K_4	−2.28	−2.28			
			$C_{i,t}$	2.14	2.14	b_4	10.92	10.92	K_5	0.30	0.30			
			$C_{d,t}$	4.19	4.19	c_1	0.30	0.30	K_6	−3.75	−3.75			
			$C_{t,d}$	v	v	c_2	v	v	K_7	0.00	0.00			
			$C_{i,d}$	1.28	1.28	c_3	0.30	0.30	K_8	−3.40	−3.40			
			$\frac{\hat{r}_{t,t}}{\hat{r}_{i,t}}$	v	v	c_4	v	v						
						c_5	v	v						
						e	2.00	2.00						
						f	3.00	3.00						

Note.—v = variable.

$c = \left(\frac{b}{4}\right)^{1.1} \log\left(\frac{B_t}{B_i} + 1\right)$. This is empirically derived and not yet based upon theory.

by the suprathreshold test light, according to Equation 17 above. The increase in test area therefore has no further effect on the discharge off fibers thereby resulting in little or no disinhibition or enhancement.

Summary

The brightness of a test-field rectangle was studied as a function of its area and the luminance of a nearby inducing-field rectangle. The area of the test field was varied by increasing its vertical dimension from 5.5′ to 33.0′ in visual angle. A binocular matching technique was used in which *S* indicated the test-field brightness by the luminance of a match-field set equal in brightness to the test field.

Two experiments were performed. In the first the match area was held at a constant maximum value. In the second the match area was allowed to vary along with the test area. The dependent variable for both experiments was match-field luminance and the independent variables, test area, test luminance, and inducing luminance. Very little if any change occurred in the test-field brightness as a function of test-area variation in either experiments. This held true for a wide range of test and inducing luminances. As inducing luminance was increased, however, test brightness became depressed. A mathematical theory fit to the data was based upon physiological inhibitive interaction among individual cone receptors in the retina.

References

Adrian, E. D. *The basis of sensation.* London: Christopher, 1928.

Barlow, H. B., FitzHugh, R., & Kuffler, S. W. Resting discharge and dark adaptation in the cat. *J. Physiol.*, 1954, **125,** 28–29.

Boynton, R. M., & Riggs, L. A. The effect of stimulus area and intensity upon the human retinal response. *J. exp. Psychol.*, 1951, **42,** 217–226.

Diamond, A. L. Foveal simultaneous brightness contrast as a function of inducing and test-field luminances. *J. exp. Psychol.*, 1953, **43,** 189–195.

Diamond, A. L. Foveal simultaneous contrast as a function of inducing-field area. *J. exp. Psychol.*, 1955, **50,** 144–152.

Diamond, A. L. A theory of depression and enhancement in the brightness response. *Psychol. Rev.*, 1960, **67,** 168–198.

Fry, G. A., & Alpern, M. The effect of a peripheral glare source upon the apparent brightness of an object. *J. Opt. Soc. Amer.*, 1953, **43,** 189–195.

Granit, R. *Receptors and sensory perception.* New Haven: Yale Univer. Press, 1955.

Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.

Hartline, H. K. The response of single optic nerve fibers of the vertebrate retina. *Amer. J. Physiol.*, 1938, **121,** 400–415.

Hartline, H. K., & Graham, C. H. Nerve impulses from single receptors in the eye. *J. cell. comp. Physiol.*, 1932, **1,** 277–295.

Hartline, H. K., & Ratliff, F. Inhibitory interaction of receptor units in the eye of *Limulus*. *J. gen. Physiol.*, 1957, **40,** 375–376.

Hartline, H. K., Wagner, H. G., & Ratliff, F. Inhibition in the eye of *Limulus*. *J. gen. Physiol.*, 1956, **39,** 651.

Heinemann, E. G. Simultaneous brightness induction as a function of inducing- and test-field luminances. *J. exp. Psychol.*, 1955, **50,** 89–96.

Leibowitz, H., Mote, F. A., & Thurlow, W. R. Simultaneous contrast as a function of separation between test and inducing fields. *J. exp. Psychol.*, 1953, **46,** 454–456.

(Received August 5, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 346–351

MEDIATED SATIATION IN VERBAL TRANSFER [1]

LEON A. JAKOBOVITS AND WALLACE E. LAMBERT

McGill University

The role of mediation in associative processes has long been recognized by psychologists (see Atherton & Washburn, 1912) despite early failures to demonstrate its influence in controlled experiments (e.g., Howe, 1893). Peters (1935) was among the first to report positive results in controlled designs, and recently, several investigators (McGehee & Schulz, 1962; Russell & Storms, 1955) showed conclusively that mediation in paired-associate learning can be demonstrated in the laboratory. In these studies the existence of mediation in verbal learning was inferred from transfer designs permitting positive generalization of acquisition gradients. The present paper represents an attempt to study mediation in verbal transfer using generalization of inhibition gradients. In other words, a design was used which, on the basis of previously established facts, should lead to facilitative effects in paired-associate learning but, because the inferred mediators are "tampered with," inhibition instead of facilitation is expected to be transferred. Since the present design produces inhibition by the direct manipulation of the mediators, it has an advantage over previous approaches in which the mediated effect remained at the implicit level.

[1] This research was supported by the Canadian Defense Research Board Grant D77-9401-10 and in part by a subvention from the Carnegie Corporation. We are grateful to Rabindra Kanungo for helpful suggestions.

METHOD

Design.—The design of the present study was parallel to those of Russell and Storms (1955) and McGehee and Schulz (1962). In Group E (experimental), *S*s learned two paired-associate lists: List 1 established A-B connections between nonsense syllables and meaningful words; List 2 was composed of A-D pairs, where the relation between B and D was such that D was the most common associate to C and the latter was the most common associate to B. The middle element, C, acts as the mediating link in the forward associative chain, $B \rightarrow C \rightarrow D$, and provides facilitation in the acquisition of the A-D list. In the present case, however, the C word was "satiated" according to a technique described by Lambert and Jakobovits (1960) who showed that continuous repetition of a word results in a decrement in the intensity of its connotative meaning. It was expected that the decreased meaningfulness of the C word would reduce its effectiveness to act as a mediator in the $B \rightarrow C \rightarrow D$ chain, reducing the facilitation effect during the subsequent acquisition of the A-D pairs.

In Group C (control), *S*s also learned two lists: List 1 established A-X associations between the same nonsense syllables but other meaningful words than those used for Group E; List 2 was composed of the same A-D pairs as used for Group E. However, no associative relation existed between the X and D words. The words were the same as those used in the Russell and Storms (1955) study where a complete description of the procedure for selecting words can be found.

The overall design is illustrated in Table 1. Group E received the Mediator Nonsatiated and the Mediator Satiated conditions. Each of these two conditions consisted of five paired associates. The A_1-B_N and A_1-D_N pairs correspond to the second half of the A-B and A-D pairs, respectively, in Table 3 of the Russell and Storms study. The A_7-B_8 and A_7-D_8 pairs correspond to the first half of the A-B and A-D pairs, respectively, in their study. The 10 A-B pairs formed List 1 in the present study, while the 10 A-D pairs formed List 2. Three different random orders of the 10 pairs

TABLE 1

ILLUSTRATION OF THE DESIGN USED IN THE EXPERIMENT

Condition	List 1	List 2	Inferred Action
Mediator Nonsatiated	A_1-B_N	A_1-D_N	$A_1 \rightarrow B_N \rightarrow C_N \rightarrow D_N$
Mediator Satiated	A_2-B_S	A_2-D_S	$A_2 \rightarrow B_S \rightarrow C_S \rightarrow D_S$
Nonmediated Control	A_1-X_1	A_1-D_N	$A_1 \rightarrow X_1 \rightarrow ?$; $A_1 \rightarrow D_N$
	A_2-X_2	A_2-D_S	$A_2 \rightarrow X_2 \rightarrow ?$; $A_2 \rightarrow D_S$

Examples of the Above Conditions in the Same Order

List 1	List 2	Inferred Action
GEX-JUSTICE	GEX-WAR	JUSTICE→PEACE; GEX→WAR
YOV-SOLDIER	YOV-NAVY	SOLDIER→ARMY; YOV→NAVY
GEX-HOUSE	GEX-WAR	HOUSE→?; GEX→WAR
YOV-CHEESE	YOV-NAVY	CHEESE→?; YOV→NAVY

Note.—The notations A_1 and A_2 refer, respectively, to the first 5 and second 5 stimulus members of the 10-item lists. Similarly, B_N or D_N and B_S or D_S refer to the first and second half of the response members.

in each list were presented in a standard memory drum at a 3:3-sec. rate with a 6-sec. intertrial interval. The instructions given were the same as those described in detail by Storms (1958). Each *S* saw List 1 for a maximum of 27 trials or until he met the criterion of three errorless repetitions, whichever came first. (All *S*s met the lesser criterion of one errorless repetition, but 10 failed to meet the criterion of three errorless repetitions within the maximum of 27 trials.) Eight minutes elapsed between the last presentation of List 1 and the first presentation of List 2. During this period, *S*s of Group E received the satiation treatment which involved the five C_S words of the Mediator Satiated condition (A_2-B_S), as well as five filler words (actually, the second half of the X words in the Russell and Storms table). First, *S* rated the 10 words (randomly mixed)

TABLE 2

NUMBER OF TRIALS TO CRITERION ON THE VARIOUS CONDITIONS FOR GROUPS E AND C

Condition	Group E		Group C		*t* between Groups
	Mean	*SD*	Mean	*SD*	
A_1-B_N and A_2-B_S (List 1)	19.80[a]	5.09	—	—	1.26
A_1-X_1 and A_2-X_2 (List 1)	—	—	17.95[b]	4.22	
A_1-D_N and A_2D_S (List 2)	10.77[a]	4.85	17.30[b]	5.16	4.45**
A_1-D_N	7.77	3.45	15.25	5.14	6.13**
A_2-D_S	10.30	4.74	14.45	5.16	2.86*

[a] The difference between these two means is significant ($t = 11.73$; $P < .001$).
[b] The difference between these two means is not significant ($t = 0.64$; *ns*).
* $P < .01$.
** $P < .001$.

on three scales of the semantic differential (pleasant-unpleasant, strong-weak, and fast-slow). Then *S* repeated aloud each word for 20 sec. before rating it again on the same three scales. Differences in intensity of ratings before and after repetitions represent the semantic satiation scores to be presented below. Following the satiation treatment, List 2 was presented and *S* learned it to a criterion of three errorless repetitions and in this case all *S*s reached the criterion in less than 27 trials.

The *S*s in Group C received the Nonmediated Control condition shown in Table 1. The five A_1-X_1 pairs and the five A_2-X_2 pairs formed List 1 which corresponds to the A-X column of the Russell and Storms table. The A_1-D_N and A_2-D_S pairs which formed List 2 were the same as those for Group E. The procedure used with Group E was duplicated here except for the fact that no satiation treatment was administered, and *S*s were engaged in neutral conversation during the 8 min. which separated the two lists.

In summary, the following comparisons can be noted: the Mediator Nonsatiated condition in the present study corresponds to the "Chained" condition of Russell and Storms (1955) or the "Mediated" group of McGehee and Schulz (1962). The Nonmediated Control condition corresponds to the "Unchained" or "Nonmediated" conditions respectively in those two studies; the Mediator Satiated condition represents the proactive interference design of the present study. The results are presented in terms of differences in acquisition scores of List 2 under the three conditions.

Subjects.—The *S*s were 50 male English-speaking cadets of the Royal Canadian Air Force enrolled in a 6-wk. training course at a base near Montreal. They were asked to volunteer for the experiment by their instructor. The testing was done individually at the training base during regular work hours. The first 30 *S*s formed Group E, the last 20, Group C.

Results

Comparison between the two groups on the number of trials required to reach the criterion of one errorless repetition (met by all *S*s) for their particular first list is presented at the top of Table 2. There is no significant difference between the two groups, indicating that they are of

TABLE 3

Comparisons between the Mediator Nonsatiated (A_1-D_N) and Mediator Satiated (A_1-D_S) Conditions of List 2 for Groups E and C

Measures	Means for Group E			Means for Group C		
	A_1-D_N	A_2-D_S	Diff.	A_1-D_N	A_2-D_S	Diff.
First 5 diff. correct responses	3.13 (0.62)	1.87 (0.62)	1.26** (1.72)	2.45 (1.07)	2.55 (1.07)	−0.10 (2.18)
Trials to 1 errorless repet.	7.77 (3.47)	10.30 (4.73)	−2.53** (3.28)	15.25 (5.14)	14.45 (5.15)	0.80 (4.05)
Total correct anticipations[a]	33.07 (15.40)	26.70 (11.83)	6.37** (6.41)	43.30 (17.10)	44.10 (15.50)	−0.80 (14.52)
Total number of intrusions[a]	3.97 (3.43)	6.30 (4.60)	−2.33* (3.50)	12.30 (6.45)	12.80 (7.65)	−0.50 (8.07)
Total number of omissions[a]	17.67 (11.43)	21.71 (13.83)	−4.04** (4.78)	30.25 (7.55)	28.95 (8.85)	1.30 (7.98)
Total correct anticipations[b]	47.90 (16.20)	40.83 (11.83)	7.07** (7.37)	56.10 (14.85)	57.25 (13.75)	−1.15 (13.19)

Note.—Numbers in parenthesis are *SD*s.
[a] To a criterion of one errorless repetition.
[b] To a criterion of three errorless repetitions or 27 trials.
* $P < .01$, two-tailed t test.
** $P < .001$, two-tailed t test.

equal learning ability. This conclusion is correct unless the materials in the two lists are not of equal difficulty which is unlikely in view of the particular selection procedure used by Russell and Storms (1955) and the fact that McGehee and Schulz (1962), using the same meaningful words but other nonsense syllables of comparable association values, found no such difference. The two groups differ significantly on the acquisition of List 2, with Group E showing a marked superiority over Group C. This finding is a replication of McGehee and Schulz and other studies in which the mediated condition was found to be superior to a nonmediated condition. The same finding is pointed up by the fact that the acquisition of List 2 by Group E is significantly faster than for List 1, whereas no such facilitation effect is noticeable for Group C (see Footnotes a and b to Table 2). Breaking down the analysis of List 2 acquisition into A_1-D_N and A_2-D_S pairs it can be seen that in both cases Group E is significantly superior to the control. This means that the predicted proactive inhibition effect of the Mediator Satiated condition (A_2-D_S) did not result in absolute negative transfer. In fact, facilitation was noticed, although significantly less than in the case of the Mediator Nonsatiated condition (A_2-D_N), as will be indicated below.

The mean semantic satiation score for Group E on all 10 words was -3.07 ($SD = 6.64$; $t = 2.49$; $P < .02$). For the five C_S words the mean was -2.27 ($SD = 3.82$; $t = 3.20$; $P < .01$); the mean for the five filler words was -0.80 ($SD = 4.09$; $t = 1.05$; *ns*). This difference in amount of semantic satiation shown on the two sets of words approaches significance ($t = 1.84$; $P < .10$), even though the product-moment correlation coefficient between the two sets is significant ($r = .43$; $P < .02$). Although there is a basis for speculating on the reason for the difference noted (e.g., the C_S words are related to the B words seen in List 1, whereas the filler words are not so related), such a discussion is not directly relevant to the purpose of this study.

Let us then turn to a comparison of the acquisition scores between the A_1-D_N and A_2-D_S pairs of List 2 for Group E. Russell and Storms (1955), making a similar comparison in their study, used two separate criteria. One of these was an analysis of the first five different correct responses made by the *S*s, with a view to determine whether the response terms of the A-D pairs, for which associative chaining was possible (here, the A_1-D_N pairs), were more easily elicited during the early trials. In the present study these were compared to the A_2-D_S pairs where the mediator was satiated. The other analysis involved subtracting the total number of correct anticipations by each *S* for the Mediator Satiated pairs (A_2-D_S) from the corresponding total for the Mediator Nonsatiated pairs (A_1-D_N). If there is inhibition during learning of the Mediator Satiated pairs (i.e., *S* has a smaller number of correct anticipations for the A_2-D_S than the A_1-D_N pairs), this difference will be positive. In view of Weitz's (1961) argument that conclusions based on verbal learning data often depend on the type of criterion measure used, it was decided to add four other measures in the present analysis and these are presented in Table 3. It can be seen that all six measures used in the comparison support the conclusion that semantic satiation of the connotative meaning of the mediator (C) increases the difficulty with which an A-D list is acquired after having established

A-B pairs. Furthermore, the proactive inhibition effect is noticeable not only as an increase in the number of trials required for acquisition, but also as an increase in both the number of intrusions and omissions made during acquisition. The same analysis for Group C is also given in Table 3. It will be remembered that Group C established A-X connections before learning A-D associations, and no satiation treatment was given. It can be seen that none of the six measures indicate differential difficulty of acquisition of the A_1-D_N and A_2-D_S pairs.

A closer examination of the data lends support to the main finding. If the A_1-B_N pairs of List 1 had been originally better learned by experimental *S*s than the A_2-B_S pairs, then the A_1-D_N pairs of List 2 would have been easier to acquire than the A_2-D_S pairs. Analysis of the total number of correct anticipations of the A_1-B_N pairs minus the corresponding number for the A_2-B_S pairs during learning of List 1 (to a criterion of one errorless repetition) yielded a mean of -2.47 ($SD = 15.51$; $t = 0.88$; *ns*). Thus, not only is there no reliable difference in the learning of A_1-B_N and A_2-B_S pairs in List 1, but also the A_1-B_N pairs were somewhat more difficult to learn as indicated by the minus value of the mean difference. Consequently, the differential difficulty of A-D pairs under the satiation and nonsatiation treatments cannot be attributed to the initial differential difficulty of A-B pairs. Also, the fact that Group C did not exhibit superiority of the A_1-D_N over the A_2-D_S pairs of List 2 (Table 3) indicates that there was no intrinsic difference between the two sets of A-D pairs. Finally, $r = .42$ ($P < .02$) between the degree of semantic satiation shown by each experimental *S* on the satiated mediators (C_S) and the extent of inhibition shown in acquisition of A_2-D_S pairs relative to the A_1-D_N pairs. This is a most interesting finding since it shows that the extent of proactive inhibition for individual *S*s caused by satiation of the mediator is related to the degree of decreased meaningfulness of the mediator itself.

Discussion

The finding that a significant facilitation effect is obtained in List 2 with Group E, but not with Group C, is a replication of the positive transfer effect of the mediation paradigm reported by several authors and will not be discussed further in detail. The specific contribution of this paper concerns the other finding reported, namely that significantly less facilitation is obtained when the mediator is satiated. Two possible mechanisms might be operating here: one is that the availability of the mediator is reduced, making the completion of the mediation sequence $B \rightarrow C \rightarrow D$ less probable; the other is that, given the assumption (see Staats, 1961) that some of the mediation reactions in the mediator C are also involved in B and D, the inhibition process might generalize to these two terms as well. As a result of either of these two processes, the subsequent acquisition of D will have been made more difficult. Furthermore, the amount of generalized inhibition or secondary extinction might be expected to be proportional to the degree to which the original stimulus word was inhibited. The significant positive correlation reported above is consistent with this expectation.

The authors view the present findings as further support for their interpretation of semantic satiation as a cognitive form of reactive inhibition having characteristics similar to extinction phenomena noted with conditioned responses (Jakobovits & Lambert, 1962a, 1962b; Lambert & Jakobovits, 1960). The present study also shows that the role of media-

tion in verbal learning can be studied from the point of view of generalization of inhibitory processes—an approach which is complementary to the positive semantic generalization paradigms used so far by previous investigators. Studies on the generalization of semantic inhibition, or, as in the present case, of mediated satiation, ought to prove useful in eliminating undesirable verbal habits, and perhaps even nonverbal habits, and may well provide a tool for the experimental manipulation of implicit verbalizations or meanings in studies of thinking and problem solving.

Summary

The present study was concerned with mediation in verbal transfer where generalized inhibition could be observed from one learning task to another. A proactive interference paradigm was arranged using the same method and material as Russell and Storms (1955) and McGehee and Schulz (1962). Assuming that mediation follows the sequence B → C → D, *Ss* first learned an A-B list, then the meaning of the inferred mediator, C, was reduced by a satiation procedure, and finally an A-D list was learned. It was shown that satiation of the mediator resulted in generalization of inhibition during learning of the A-D list. The results were discussed in the light of the existing theoretical models for mediated generalization.

References

Atherton, M. V., & Washburn, M. F. Mediate association studied by the method of inhibiting associations: An instance of the effect of "Aufgabe." *Amer. J. Psychol.*, 1912, **23**, 101–109.

Howe, H. C. "Mediate" association. *Amer. J. Psychol.*, 1893, **6**, 239–241.

Jakobovits, L. A., & Lambert, W. E. Semantic satiation in an addition task. *Canad. J. Psychol.*, 1962, **62**, 112–119. (a)

Jakobovits, L. A., & Lambert, W. E. Semantic satiation among bilinguals. *J. exp. Psychol.*, 1962, **62**, 576–582. (b)

Lambert, W. E., & Jakobovits, L. A. Verbal satiation and changes in the intensity of meaning. *J. exp. Psychol.*, 1960, **60**, 376–383.

McGehee, N. E., & Schulz, R. W. Mediation in paired-associate learning. *J. exp. psychol.*, 1962, **62**, 565–570.

Peters, H. N. Mediate association. *J. exp. Psychol.*, 1935, **18**, 20–48.

Russell, W. A., & Storms, L. H. Implicit verbal chaining in paired-associate learning. *J. exp. Psychol.*, 1955, **49**, 287–293.

Staats, A. W. Verbal habit-families, concepts, and the operant conditioning of word classes. *Psychol. Rev.*, 1961, **68**, 190–204.

Storms, L. H. Apparent backward association: A situational effect. *J. exp. Psychol.*, 1958, **55**, 390–395.

Weitz, J. Criteria for criteria. *Amer. Psychologist*, 1961, **16**, 228–231.

(Received August 9, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 352–354

WORK DECREMENT AND REMINISCENCE IN PIGEON OPERANT RESPONDING[1]

C. ALAN BONEAU AND SEYMOUR AXELROD

Duke University

The present study was prompted by the incidental observation that pigeons in a Skinner box seemed to exhibit some of the phenomena of human motor learning (see, e.g., Kimble, 1949). Specifically, after the pigeons had been trained to peck a key with a pattern projected on it, and not to peck when the key was blank (a relatively distributed work schedule), response rate declined over a series of successive positive periods (work decrement during relatively massed practice) and then increased markedly after a negative period in which response rate was low or zero (reminiscence). The experiment described here was a more systematic exploration of these phenomena; additionally, the question was asked whether the reminiscence effect is attributable merely to the rest during the negative period, or to inhibitory properties acquired by the negative stimulus during discrimination training.

METHOD

The *S*s were 12 naive young-adult white Carneau pigeons, maintained at 75% ad lib. weight by manipulation of food intake. The *S*s were first key trained and then given 6 days of training on a 60-sec. variable interval (VI) schedule for 30 1-min. periods per day. Following this, *S*s were trained for 4 days, using a 60-sec. VI schedule for positive periods and no reinforcement for negative periods, to peck at a disk on which a pattern (circle with a gap at the top) was projected, and to withhold response when the disk was illuminated but blank. Response periods lasted 60 sec., and alternated with 10-sec. off periods during which the box and key were unilluminated. Thirty periods per day were run, 15 positive and 15 negative, in random sequence with the restriction that no more than 2 positives or negatives were presented successively. For the next 4 (postdiscrimination test) days, 40 periods/day were run. During Periods 9, 18, 27, and 36 the key was either (D) dark (a rest period) or (L) lighted but blank (the negative stimulus) and reinforcement was withheld; during the remaining 36 periods, the positive stimulus was present and the 60-sec. VI schedule was maintained. The temporal positions of the dark and light periods were systematically assigned so as to distribute sequence effects.

RESULTS AND DISCUSSION

In Fig. 1, average rates of responding for the 12 *S*s are plotted trial by trial for selected days of the training procedure. The first panel shows the relatively stable rate of responding on Day 6 (last) of VI training. The second and third panels show the rates to both the positive and negative stimuli for Day 1 and Day 4 (last) of discrimination training. As expected, there was a considerable warm-up effect at the beginning of each day.

The fourth panel of Fig. 1 shows the average rate per trial on the first day of the postdiscrimination training. Of interest are the high rates on the first trials of each block and the gradual decrease in rate until the interpolated blackout period or negative stimulus (Trials 9, 18, 27, and 36). The decrease is quite orderly. Individual *S*s showed a steady trial-by-trial decrease in rate, occasionally with no inversions. Least squares lines fitted to the last seven points of

[1] The authors are indebted to Norman Guttman for the use of his apparatus. David Wells assisted in running *S*s and in analyzing data.

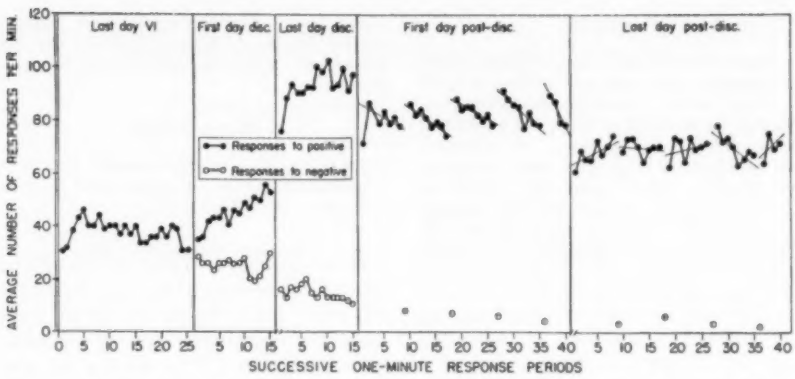

FIG. 1. Rate of responding (pecks per minute) on selected days of the training procedure.

the first group of eight trials (to compensate for warm up), and to all points for the other blocks of trials, are also shown, and are the basis for the determination of reminiscence scores, defined as the difference between the extrapolated ninth-trial performance in one block and the fitted first-trial performance in the next.

Two differences in performance between the first day and the last day of postdiscrimination (fifth panel) should be noted. First, the orderly reminiscence phenomenon occurring on the first day has disappeared by the fourth day. Secondly, there is a change in level of response. The average amount of reminiscence per bird on Day 1 of postdiscrimination, when tested against the null-hypothesized zero, yields a t of 4.19, significant for 11 df at the .01 level. Response rate declined monotonically between the last day of discrimination and the last day of postdiscrimination ($F = 6.10$, $df = 4/44$, $P < .01$).

As outlined in the procedure, half of the interpolated 60-sec. rest periods in postdiscrimination were periods in which the negative stimulus was presented; the other half were periods in which the key was not lighted. The average difference in reminiscence over all 4 days following these two kinds of interpolated periods was .48 responses per min. This difference results in a t of .09, evidence consistent with the hypothesis that the reminiscence is due only to a low level of responding (rest) during the interpolated periods.

The experiment demonstrates the occurrence in the Skinner box of a number of the phenomena usually associated with human motor learning under conditions of massed and distributed practice. The general superiority of distributed practice is evident during the latter stages of discrimination training when *Ss* are relatively inactive during the negative periods. In the postdiscrimination periods occur the phenomena of work decrement (when the work periods are relatively massed) and reminiscence following a rest. In addition, the average performance under the massed practice regime of the postdiscrimination period progressively decreases relative to distributed practice performance. Warm-up decrement is not especially evident except at the beginning of each day's performance. A final similarity to human motor performance is the decrease in the amount of reminiscence as massed practice training proceeds.

Summary

Pigeons were trained in a Skinner box to discriminate a key with and without a figure projected upon it and exhibited rapid responding to the positive stimulus and little or no responding to the negative. No more than two positives or negatives were presented successively. The *S*s were then given sequences of eight positives followed by one negative. The resulting performance exhibited phenomena characteristic of human motor learning performance under massed practice (warm-up decrement, temporary and permanent work decrement, and reminiscence).

Reference

Kimble, G. A. An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.*, 1949, **39**, 15–23.

(Received August 11, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 355–363

STUDIES OF DISTRIBUTED PRACTICE: XXII. SOME CONDITIONS WHICH ENHANCE RETENTION [1]

BENTON J. UNDERWOOD, GEOFFREY KEPPEL

Northwestern University

AND RUDOLPH W. SCHULZ

University of Iowa

The accumulated evidence indicates that in verbal learning some relatively high level of interference must be present before differences between massed practice (MP) and distributed practice (DP) will be observed. This law holds for differences in acquisition under MP and DP as well as for differences in retention following learning by MP and DP. However, the general statement relating interference and DP effects must be accompanied by subsidiary statements which relate particular DP phenomena to particular loci of the interference. A recent study (Underwood & Schulz, 1961) indicated that DP will facilitate acquisition only when interference occurs in acquiring or integrating response terms as such; DP will have no influence on retention when the interference is of this nature. On the other hand, when interference occurs in associating a response term to a stimulus term, DP may enhance retention but will have little effect on learning. It is this second subsidiary law with which the present experiments are concerned since these experiments explore further the conditions associated with better retention following DP than following MP.

In the study noted above, *S*s learned four paired-associate lists in which the stimulus terms were nonsense syllables and the response terms were two-syllable adjectives. Because of the similarity across lists among the syllables, interference in associating the response terms with the stimulus terms increased as the number of lists increased. Each successive list required the learning of new responses to stimuli which were very similar to stimuli in previous lists. This similarity may be maximized by using the *same* stimuli in all lists in which case the four lists would be symbolized as A-B, A-C, A-D, and A-E. In line with the second subsidiary law, if DP is given on A-E, learning would not be facilitated but retention would be; DP appears to reduce proactive inhibition (PI) in retention when interference is of this type. Extinction and spontaneous recovery were postulated to be the basic mechanisms which reduced PI in the situation described above where interlist stimulus similarity was high. More particularly, it was assumed that in learning A-E the associations previously acquired (B, C, and D) to A were extinguished. Moreover, when A-E learning was by DP, it was assumed that the rest intervals allowed for some spontaneous recovery of the previously acquired associations; hence, subsequent learning trials would result in the re-extinction of these associations. Finally, it was assumed that these successive extinction-recovery cycles lead to a more permanent extinction of interfering tendencies under DP than under MP,

[1] This work was done under Contract Nonr-1228 (15), Project NR 154-057, between Northwestern University and the Office of Naval Research.

thus reducing PI in retention. To say that interfering tendencies are more permanently extinguished under DP than under MP is to say only that recovery will be slower and will never attain as high a level under DP as under MP.

The major purpose of the present experiments is to attempt to identify the associations which, according to this theory, are more permanently extinguished during DP than during MP. In an A-B, A-C paradigm it is possible to identify two different sets of associations which may be extinguished in learning A-C. The most obvious set of associations which might be extinguished are the A-B associations. As *S* learns C to A, the A-B association may become extinguished. These associations will hereafter be referred to as specific S-R associations. The response terms identify a second set of associations. The B response terms may be learned without also being associated with their particular verbal stimuli in the list. That is, response learning may occur independently of specific S-R learning. For this response learning the stimulus may be the general environmental situation or, as it will be called here, the stimulus may be a contextual one. In learning the A-C list it is possible that the associations between context stimuli and the B responses become extinguished and are replaced by the C responses. Thus, we may identify at least two kinds of associations which may be extinguished in learning A-C following the learning of A-B. The experimental question asked is whether the extinction of both types of associations is involved in the better retention following DP than following MP, or whether the effect can be attributed to one or the other class of associations. That extinction of both classes of associations may occur seems to have been demonstrated by Barnes (1960). However, the relative permanence of the extinction of the two classes of associations and the relative recovery rates are unknown so that it is quite possible that the facilitating effect of DP on retention may be tied to the extinction of one class or the other rather than to the joint extinction of both classes.

In order to separate the effects of extinguishing the two classes of associations (specific S-R and contextual associations) a second paradigm is needed. This paradigm is known as the A-B re-paired paradigm in which on each successive list the same stimulus and response terms are used but the particular pairings differ from list to list. A study of this paradigm will show that in learning a second (or third, or fourth) list there are specific S-R associations which may be extinguished. However, since the same responses are used in each list, no extinction of contextual associations can occur, with the result that the contextual stimuli should remain associated for all lists. Thus, when the A-B, A-C paradigm is posed against the A-B re-paired paradigm it is seen that both have specific S-R associations which may be extinguished whereas only the A-B, A-C paradigm has contextual associations which may be extinguished. With these differences in the two paradigms in mind, the experimental situation to be used may be considered.

The *S*s will learn four lists. In one case these lists will be A-B, A-C, A-D, A-E. In the other case there will be four A-B re-paired lists. To minimize the awkwardness in referring to these two paradigms the first will hereafter be called RD (responses different in each list) and the second RS (responses same in all lists). For both

paradigms DP is introduced in the learning of the fourth lists and the effect of DP on retention measured by having other groups in which the fourth list is learned by MP.

The results to be obtained in retention following MP and DP of List 4 for these two paradigms may demonstrate the need to limit application of the recovery-extinction theory to particular kinds of associations, i.e., either to specific S-R associations or to contextual associations. This is to say that the empirical results may be used to refine the gross theory. Three different possible empirical effects of DP on retention of List 4 for the two paradigms may now be stated along with the theoretical inferences to be drawn given each effect. (*a*) Effects of DP may be positive and equal for the RD and RS paradigms. The conclusion would be that extinction of specific S-R associations is responsible for the effect. (*b*) Effects may be positive for both paradigms but greater for RD than for RS. The conclusion would be that extinction of both specific S-R associations and contextual associations is responsible for the effect. (*c*) Effects may be positive for the RD paradigm with no effect for the RS paradigm. The conclusion would be that extinction of contextual associations is responsible for the effect.

Experiments 1 and 2 were directed specifically toward testing the above notions. Experiment 3, which deals only with the RD paradigm, was designed to investigate the influence of a different level of learning than used in Exp. 1 and 2 and to make some initial determinations of differences in length of the DP interval.

Method

The three experiments have many procedural details in common. Therefore, Exp. 1 will be described in detail followed by a description of the changes introduced for Exp. 2 and 3.

Experiment 1

Design.—Two parallel sets of conditions were used, one set being based on lists forming the RD paradigm and one set based on lists forming the RS paradigm. For each paradigm four lists were learned by each *S*. For the first three lists all *S*s within a paradigm were treated alike since the learning of these lists was imposed merely to build up interlist interference. For List 4, half the *S*s under each paradigm learned by DP (1-min. intertrial interval) and half by MP (4-sec. intertrial interval). Twenty-four hours after learning List 4, recall and relearning measures were obtained on it.

Lists.—Each list consisted of eight paired associates. The stimulus terms were nonsense syllables, the response terms two-syllable adjectives. For both paradigms only eight syllables are needed as stimuli for the four lists. The syllables used are those listed in Table 1 under List 1 of the previous experiment (Underwood & Schulz, 1961, p. 229). For the RS paradigm only eight different response terms are needed for the four lists. These terms were the adjectives given for List 1 of Table 1 in the previous study. To construct the four lists the stimulus and response terms were simply re-paired randomly with the restriction that a given pairing occur in only one list. For the RD paradigm, eight different adjectives were used for each list, these four sets being those in Table 1 of the aforementioned study. List 4 was exactly the same for *S*s under both paradigms, this list being the List 1 in Underwood and Schulz (1961) except, of course, in the present study the adjectives were response terms and the syllables were stimulus terms. The sequence of Lists 1–3 was the same for all *S*s within a paradigm, i.e., they were *not* counterbalanced.

Procedure.—Lists were presented at a 2:2-sec. rate with anticipation learning throughout. Each list for *S*s in the RD paradigm was presented for 12 anticipation trials with *S* instructed to give as many correct anticipations as possible on each trial. Four different orders of items were used for each list with each order being used as a start order about equally often. Lists 1 and 2 were presented on Day 1; Lists 3 and 4 on Day 2, and 5 relearning trials were given on List 4 on Day 3, 24 hr. after Day 2. Relearning was by MP

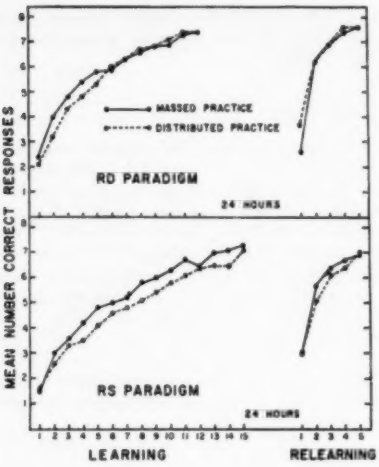

FIG. 1. Learning and relearning curves for Exp. 1.

for all groups. Approximately 30 sec. elapsed between the learning of the two lists on a given day. The DP intervals on List 4 were introduced after each trial and were filled with symbol cancellation. The procedures for *Ss* in the RS paradigm were exactly the same as for those outlined above except that 15 anticipation trials were given on each list. Since the RS paradigm produces more interference than the RD paradigm, 15 trials were used for the RS paradigm (as compared with 12 for the RD) in order to make the level of learning under the two paradigms roughly comparable.

The *Ss* were all college students, and, while all were not naive to verbal-learning experiments, none had previously learned syllable-adjective pairs. For the RD paradigm 32 *Ss* learned List 4 by DP and 32 by MP. For the RS paradigm there were 30 *Ss* in each group.

Experiment 2

As in Exp. 1, both paradigms were studied. The changes from Exp. 1 were as follows: (*a*) Lists 1–3 were given on Day 1, List 4 was given on Day 2, and List 4 was relearned for 5 trials on Day 3, 24 hr. after Day 2. (*b*) Under both paradigms 12 anticipation trials were given on all lists. (*c*) MP was 4 sec. between trials while DP consisted of a 3-min. interval after Trials 1, 3, 5, 7, 9, and 11. (*d*) All four groups (MP and DP groups for each paradigm) had 32 *Ss*.

Experiment 3

In Exp. 3 only the RD paradigm was studied. All four lists were presented during a single session but each was presented for only 5 anticipation trials. Three groups of 25 *Ss* each were used, the groups being differentiated only by the intertrial interval on List 4. One group had the usual MP (4-sec. intertrial interval), a second had 60 sec. between trials, and a third had 180 sec. between trials. Again List 4 was given 5 relearning trials 24 hr. after original learning.

RESULTS

Experiment 1.—The equivalence of the learning ability of the two groups of *Ss* within each paradigm can be gauged by the learning scores for Lists 1–3 since the conditions were identical for all *Ss* for these lists. For the RD paradigm, the mean total correct anticipations for Lists 1–3 in order for the MP *Ss* were 64.09, 64.00, and 72.31. For the DP *Ss* the corresponding values were 63.38, 63.53, and 73.03. For the RS paradigm the values for the MP *Ss* were 79.57, 80.67, and 76.73 and for the DP *Ss*, 78.40, 80.53, and 75.23.

The acquisition and relearning curves for *Ss* in both paradigms for List 4 are shown in Fig. 1. For both paradigms learning is somewhat slower under DP than under MP; this is true throughout the 15 trials for RS lists and for the initial part of learning for the RD lists. This finding confirms

that of an earlier study (Underwood & Schulz, 1961) and also conforms to the extinction theory in that the DP interval should allow interfering associations to recover, thus impeding acquisition.

Looking next at the recall trial it can be seen that for the RS paradigm recall is almost identical for both MP and DP. The slight difference in the relearning curves in favor of MP probably reflects the small difference present in original learning. It must be concluded that DP does not facilitate retention of lists forming the RS paradigm. For the RD paradigm, however, recall is better under DP than under MP. The exact values are 2.59 for MP and 3.66 for DP ($t = 2.28$, $.05 > P > .02$). If loss scores are used as described in the previous study (Underwood & Schulz, 1961, p. 231–232), in order to adjust for any differences in original learning, the evaluation does not change ($t = 2.32$). It is concluded that DP in the RD paradigm facilitates recall after 24 hr. The effect, however, is very transitory since no difference in the subsequent four relearning trials is evident.

Certain expectations concerning overt errors follow from the extinction hypothesis. First, more intrusions from previous lists should occur in learning List 4 under DP than under MP. In the RS paradigm, the definition of an intrusion is somewhat ambiguous since the same responses occur in all lists. However, the nearest approximation would be to consider only responses given to a stimulus with which that response had been paired in previous lists. Approximately two-thirds of all errors made in learning List 4 were of this nature. And, while more of these (a total of 387) occurred under DP than under MP (347), statistically speaking, the difference is far from significant. For the RD paradigm, 13 responses from List 1–3 occurred in learning List 4 under DP, and 7 under MP. While this difference is in accordance with the theory, the numbers are so small that little should be made of the effect.

A second expectation from the theory is that fewer intrusions should occur in relearning following DP than following MP. For the RS paradigm 132 intrusions (as per the definition given above) occurred in relearning following MP and 138 following DP. Clearly, there is no evidence in this paradigm of a more permanent extinction occurring under DP than under MP. For the RD paradigm, 44 intrusions occurred during relearning following MP and 24 following DP. These data are in conformance with the theory.

Experiment 2.—In this experiment, Lists 1–3 were learned on Day 1 with List 4 being learned on Day 2, and the DP interval for List 4 was 3 min. after every other trial.

In learning Lists 1–3 the mean total correct responses for the MP *Ss* in the RD paradigm were 74.59, 69.53, and 74.13. For the DP *Ss* the corresponding values were 69.78, 64.34, and 73.13. The MP *Ss* learning the RS lists showed means of 68.38, 60.97, and 67.00, while the comparable values for the DP *Ss* were 69.13, 56.56, and 64.25. Thus, the two groups within each paradigm are fairly comparable in learning ability.

The performances in learning and relearning List 4 are plotted in Fig. 2. Again the *Ss* learning RD List 4 under DP show inferior performance to those learning under MP. For the RS paradigm, the picture is a little different. After nearly every rest interval the DP *Ss* show inferior performance to the MP *Ss* but on the immediately succeeding trial the per-

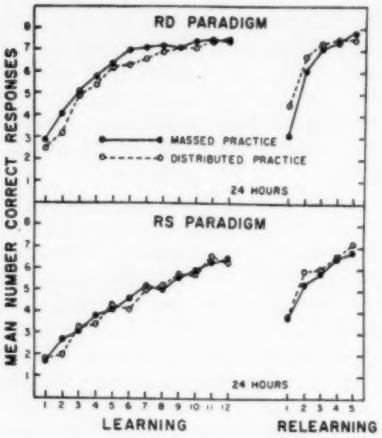

FIG. 2. Learning and relearning curves for Exp. 2.

formance of the two groups is essentially equivalent. Thus, while the rest interval appears to inhibit performance, the overall learning rate does not appear to be seriously impaired.

The recall of the DP *Ss* under the RD paradigm is again superior to the MP *Ss*. The mean values are 3.13 for MP *Ss* and 4.53 for the DP *Ss* ($t = 3.11$, $P < .01$). Clearly, DP facilitates recall of the fourth list for this paradigm. The DP *Ss* are also superior to the MP *Ss* for the first four trials of relearning. However, the difference across the five trials in terms of mean total correct responses (1.85) is not significant ($t = 1.37$).

For *Ss* learning List 4 of the RS paradigm no difference of consequence is apparent in recall and relearning. While recall favors the DP *Ss*, the difference of .11 items gives a t of only .22. Thus, as in Exp. 1, it is seen that recall of List 4 is facilitated in the RD paradigm and no appreciable effect is noted for the RS paradigm.

No difference of consequence in number of intrusions in learning under MP and DP for the RS paradigm was noted; the same was true for relearning. For the RD paradigm 13 intrusions were recorded for the MP *Ss* learning List 4 and 9 for the DP *Ss*. In this case the difference is in the opposite direction from the difference found in Exp. 1 and is not in accordance with extinction theory. For relearning, however, MP *Ss* made 55 intrusions and DP *Ss*, 28, which, as in Exp. 1, is in line with expectations from the theory.

Experiment 3.—In this experiment only the RD paradigm was used. Lists 1–4 were presented for five acquisition trials each, with three subgroups of 25 *Ss* each having 4, 60, or 180 sec. as the intertrial interval on List 4. The mean total correct responses given on Lists 1–3 by the 4-sec. *Ss* were 17.76, 14.92, and 19.52, respectively. The comparable scores for the 60-sec. *Ss* were 17.80, 14.92, and 20.28, and for the 180-sec. *Ss*, 17.88, 14.96, and 21.32.

The learning and relearning curves for List 4 for the three groups are shown in Fig. 3. Again it is to be noted that during learning the performance of the two DP groups is inferior to that of the MP group. Both DP curves are consistently lower than the MP curve throughout the five trials. Again, however, the dif-

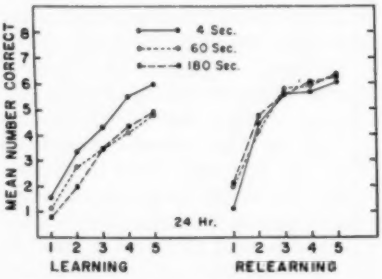

FIG. 3. Learning and relearning curves for Exp. 3.

ferences are just short of significance at $P = .05$ ($F = 2.91$, $df = 2/72$). Nevertheless, in view of the fact that for this paradigm DP has been consistently inferior to MP in learning in all three experiments, the inhibitory effect may be taken to be reliable.

The mean number of responses correctly recalled after 24 hr. were 1.20, 2.04, and 2.12 for 4, 60, and 180 sec., respectively ($F = 1.95$, $df = 2/72$, $P > .05$). However, it seems clear that adjustments must be made to compensate for differences in level of learning originally attained by the three groups. When this adjustment is made by the method noted in the earlier paper (Underwood & Schulz, 1961), the F becomes 21.40 ($df = 2/72$, $P < .001$). Therefore, it may again be concluded that for the RD paradigm distributed practice reduces PI in retention.

In the results of Exp. 1 and 2 it was noted that the facilitation in retention for the RD paradigm produced by DP was largely limited to the first recall trial. In the present data the effect is much less transitory; in spite of the fact that MP *S*s were appreciably better in learning, all groups are about equal in relearning. This effect of DP on relearning may be estimated by subtracting for each *S* the number of correct responses given on the five learning trials from the number of correct responses given on the five relearning trials. The means for these differences were 2.16, 7.84, and 9.20 for the 4-, 60-, and 180-sec. *S*s, respectively. The F (12.60) is beyond the .01 level. The potency of the PI effects on the MP *S*s can be further described by noting that of the 25 *S*s, 11 showed performance in relearning which was inferior to that shown during learning. By the same token the reduction in PI effects produced by DP can be seen by noting that only one of the 60-sec. *S*s performed more poorly in relearning than in learning and none of the 180-sec. *S*s performed so.

In learning List 4 the 4-sec. *S*s gave 7 intrusions, i.e., responses which were appropriate for an earlier list. The 60-sec. *S*s gave 32 such responses and the 180-sec. *S*s gave 11. In the 4-sec. condition all 7 intrusions were given by a single *S*, whereas 16 *S*s gave intrusions under the 60-sec. condition and 4 under the 180-sec. condition. Thus, while more intrusions were made under the DP conditions than under the MP condition, the fact that the 60-sec. *S*s gave more intrusions than did the 180-sec. *S*s would not be anticipated by the extinction hypothesis.

The intrusions during relearning totaled 34, 33, and 25, for the 4-, 60-, and 180-sec. *S*s, respectively. The theory predicts more intrusions for MP *S*s than for DP *S*s. While the values indicate this to be the case the differences are so small that they do not allow any strong support for this aspect of the theory.

Discussion

The results of Exp. 1 and 2 will be evaluated first. The basic findings were as follows. When *S*s learn four lists with the RS paradigm, DP on List 4 does not facilitate the 24-hr. retention of this list. But, when *S*s learn four lists with the RD paradigm, the retention of List 4 is facilitated by learning the list under DP conditions. In terms of the argument advanced in the introduction, the implication of these facts is that retention of a list learned by DP is facilitated only when the paradigm allows extinction of contextual associations (associations between the general environment and the response terms). Both paradigms studied allow, presumably, extinction of specific S-R associations but only the RD paradigm allows for extinction

of contextual associations. Thus, the notion that DP, by allowing for successive recovery-extinction cycles, leads to a more permanent extinction of interfering tendencies appears to be supported for contextual associations only.

The question may be raised as to why the successive-extinction hypothesis is not supported in the case of specific S-R associations in view of the fact that the extinction of such associations seems to have been demonstrated (Barnes, 1960; Barnes & Underwood, 1959). One possibility is that recovery of such associations occurs very slowly; thus, the short DP intervals used here may not in fact allow for recovery-extinction cycles. And in fact there is evidence (e.g., Briggs, 1954) that recovery of such associations is indeed very slow. A barrier to the acceptance of this position is the fact that in learning List 4 in the RS paradigm there was evidence for *some* process which was impeding performance under DP. It is reasonable to think that this could represent the recovery of error tendencies of some sort. Yet, it is possible that the interference which appears to increase with DP of List 4 in both paradigms is not representative nor symptomatic of the process which produces the better retention following DP of List 4 under the RD paradigm. The inconsistency from experiment to experiment of the differences in intrusion frequency between MP and DP on List 4 might argue for such a position. Therefore, although slower learning of List 4 by DP than by MP is consistent with an extinction hypothesis, the slower learning need not necessarily be taken to mean that better retention occurs as a consequence.

Empirically speaking, there is a relatively simple principle which may be stated which will summarize the situations in which DP may be expected to facilitate retention. As noted in the introduction, to expect any facilitation in retention by DP during learning requires first that appreciable interlist interference be present. Given this situation, the principle is that whenever the response terms of the previously learned associations producing the interference are *not* present in the list being learned, DP will facilitate the retention of the list. This empirical induction not only conforms to the paradigms and findings of the present experiments but also will handle other findings. For example, when there is high intralist similarity among syllables within a list, no effect on retention is noted if the list is learned by DP (Underwood & Richardson, 1958). The interference falling on a given association in this situation is caused by letter sequences which are appropriate or correct for other associations in the list; that is, the response units of the associations producing the interference are correct for other items in the list. The same situation holds for the RS paradigm in the present studies. To account for this empirical generalization—the generalization that DP will facilitate retention only when the response terms for the associations producing the interference are not in the list being learned—we have used the notion of recovery-extinction cycles leading to a more permanent extinction of contextual associations.

Why these contextual associations "behave" differently than other associations (e.g., specific S-R associations) is not known. Of course, it should be remembered that the present experiments have only scratched the surface in terms of studying what appear to be the relevant variables and their interaction. Degree of learning of the interfering lists, degree of learning of the list being interfered with and to be recalled, length of DP interval, and length of retention interval, all should be pertinent variables. Clearly they are pertinent in terms of the extinction-recovery theory but they undoubtedly would also be judged to be of importance by any careful empirical analysis.

The final point of discussion concerns the results of Exp. 3. In this experiment the RD paradigm was used and, in conformance with the results of Exp. 1 and 2 with this paradigm, facilitation in retention of List 4 was observed following learning by DP. The effects on retention were rather substantial in that

relearning was enhanced following DP. Clearly, PI was reduced by DP. It had been expected that the 3-min. DP interval would produce more facilitation than would the 1-min. DP interval. There was no strong evidence of differences in the effect of these two conditions. It is true that relearning was a little better following a 3-min. DP interval than following the 1-min. interval. Also, analyses of particular items showed that items having the greatest number of correct anticipations at the end of learning of List 4 produced higher recall following 3-min. DP than following 1-min. DP. Still, the effects were not as great as anticipated. In terms of the extinction-recovery notion, it might be argued that 1 min. was sufficient for the recovery of all interfering tendencies for the relatively low degree of learning of the interfering lists used. If this is true, a more direct relationship between retention and length of intertrial interval would be expected if the degree of learning of the interfering associations were higher. A comparison of the results for the RD paradigm of Exp. 1 and 2 would tend to support this notion. In these experiments the degree of learning of the interfering associations was much stronger than in Exp. 3. The 3-min. interval of Exp. 2 gave somewhat greater facilitation in retention than did the 60-sec. DP interval used in Exp. 1. However, this is by no means a clear test since other factors also varied between the two experiments. In any event, some conviction is held for the notion that in most situations a direct relationship between length of intertrial interval and facilitation in retention will be found.

Summary

Three experiments were performed to study retention following massed (MP) and distributed practice (DP) when interlist interference was high. In Exp. 1 and 2 half the *Ss* learned four lists of paired associates (syllable-adjective pairs) in which the stimuli were identical across all lists but with the responses different in each list (RD paradigm). The other *Ss* learned four lists in which the stimuli and responses were identical across all lists but with different pairings for each list (RS paradigm). Distributed practice was introduced in learning List 4 with retention of this list measured after 24 hr. In Exp. 1, the DP interval was 60 sec. between each trial while in Exp. 2 the interval was 3 min. after every other trial; MP consisted of 4 sec. between trials for both experiments.

The results were essentially the same for Exp. 1 and 2; DP facilitated recall of List 4 only for the RD paradigm. These findings indicate that DP will facilitate retention only when the response terms of previously acquired associations producing the interference are not present in the list being learned by DP. Theoretically, the results imply that DP allows for a more permanent extinction of contextual associations but does not influence specific S-R associations.

Experiment 3 used only the RD paradigm with a low degree of learning of all four lists and with intertrial intervals of 4, 60, and 180 sec. on List 4. Distributed practice markedly facilitated both recall and relearning of List 4. However, no appreciable difference was noted in the results for the 60- and 180-sec. intervals.

References

Barnes, J. M. "Fate" revisited. Unpublished doctoral dissertation, Northwestern University, 1960.

Barnes, J. M., & Underwood, B. J. "Fate" of first-list associations in transfer theory. *J. exp. Psychol.*, 1959, **58**, 97–105.

Briggs, G. E. Acquisition, extinction, and recovery functions in retroactive inhibition. *J. exp. Psychol.*, 1954, **47**, 285–293.

Underwood, B. J., & Richardson, J. Studies of distributed practice: XVIII. The influence of meaningfulness and intralist similarity of serial nonsense lists. *J. exp. Psychol.*, 1958, **56**, 213–219.

Underwood, B. J., & Schulz, R. W. Studies of distributed practice: XX. Sources of interference associated with differences in learning and retention. *J. exp. Psychol.*, 1961, **61**, 228–235.

(Received August 25, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 364–372

THE ROLE OF RESPONSE SIMILARITY IN PROACTIVE INHIBITION[1]

KENT M. DALLETT[2]

University of California

It has generally been supposed that interlist response similarity is an important determinant of interlist interference. However, while variations in response similarity have several times been shown to affect retroactive inhibition (RI) (Slamecka & Ceraso, 1960), effects of interlist similarity on proactive inhibition in retention (PI) have not yet been demonstrated. Young (1955) failed to find significant differences in PI as a function of response similarity, although he did find the expected differences in RI. Similar findings with respect to PI were reported by Morgan and Underwood (1950), the only significant PI in their study resulting from the use of identical stimuli and dissimilar responses. In each of the studies mentioned, however, there was an insignificant trend suggesting an effect similar to that generally found in RI; i.e., recall improved as response similarity increased. This trend was directly opposed to the differences predicted by Young (1955), who supposed that as response similarity increased, greater generalized strength would be added to List 1, increasing its potential for interfering with List 2 recall.

Recently, Barnes and Underwood (1959) have proposed separate mechanisms for the learning of similar and dissimilar responses in List 2. When the responses in List 2 are dissimilar to those on List 1 (with identical stimuli: A-B, A-C), List 1 responses are unlearned during List 2 learning; while with similar responses, it is suggested that *S* makes use of the List 1 response as a mediator. Thus, in learning A-B′ after A-B, *S* is assumed to be learning A-(B)-B′, although *S*s report the dropping-out of the mediator as learning progresses. Accepting Barnes and Underwood's analysis, and assuming that PI results from the response competition produced when unlearned List 1 responses spontaneously recover in strength, one might be led to expect considerable PI in the A-B, A-C paradigm, but little or no PI in the case of A-B, A-B′ (Postman, 1961). This analysis can readily be extended to intermediate degrees of response similarity by assuming that as response similarity increases from dissimilarity, unlearning gradually gives way to mediation. In this way, the data of Morgan and Underwood (1950) and of Young (1955) can be explained.

However, we do not know that the unlearning-and-recovery sequence is the only mechanism involved in PI. It is possible, for example, that similar responses may lead to loss of differentiation (Underwood, 1945) with the passage of time. Thus, *S*s may recall both responses, but may not remember which response belongs in List 2. One might, on the other

[1] This paper is based upon a dissertation submitted in partial fulfillment of the requirements for the PhD degree at the University of California. The author is grateful to Leo Postman for his enthusiastic advice and encouragement at all stages of the research. At the time, the author was a National Institute of Mental Health Predoctoral Research Fellow, on Fellowship MF 12-379.

[2] Now at University of California, Los Angeles.

hand, expect proactive facilitation if the List 1 mediator protects List 2 responses from the extraexperimental interference to which a control group may be subject. These possibilities hold for delayed recall only, little PI being expected at short intervals in any case. Indeed, since PI appears to increase with time, it would seem advisable to investigate the effects of response similarity on PI using retention intervals considerably longer than the 20-min. retention intervals employed by both Morgan and Underwood (1950), and Young (1955).

In the experiment to be reported, independent groups of *Ss* learned two lists with identical stimuli (S_I) and highly similar (S_IR_1), less similar (S_IR_2) or dissimilar (S_IR_N) responses, different *Ss* being tested for recall of List 2 either 30 sec. or 48 hr. after learning. In addition, 30-sec. and 48-hr. groups were tested under each of two control conditions: (*a*) the standard PI control group, having no first list (No PL), and (*b*) a group having dissimilar stimuli and responses on the two lists (S_NR_N). This second control condition served as a control for warm-up and learning-to-learn effects in the learning of List 2, and also was intended to sample the low-similarity end of the continuum of stimulus similarity.

METHOD

Materials and apparatus.—The words used as responses were chosen from Haagen's (1949) norms. Subjects in all conditions learned the same List 2. Corresponding to a given word on List 2, words were chosen for three different first lists so as to be highly similar (R_1), less highly similar (R_2), or dissimilar (R_N) to the List 2 word. The R_1 and R_2 lists had mean Haagen similarities of 1.40 and 3.36, respectively. The R_N list was made up of Haagen words not in the same category as the corresponding List 2 response. In addition, an attempt was made to match corresponding words for familiarity using Haagen's norms, and for frequency of usage using the Lorge magazine count (Thorndike & Lorge, 1944). Thus, the three sets of List 1 responses, as well as the List 2, or "standard" set, were matched both in mean frequency and familiarity and in frequency and familiarity of corresponding words. The responses are presented in Table 1.

TABLE 1

RESPONSES USED IN THE EXPERIMENT

List 2	List 1		
	R_1	R_2	R_N
AGILE	NIMBLE	ALERT	UNKIND
BELOVED	CHERISHED	VALUED	PETTY
COMPLETE	ENTIRE	PERFECT	HEAVY
CRAFTY	CUNNING	STEALTHY	PRIOR
DISTANT	REMOTE	FURTHER	SPOKEN
DECEASED	LIFELESS	EXTINCT	OBSCENE
FRUITFUL	FERTILE	PREGNANT	IMPURE
HAUGHTY	SNOBBISH	SCORNFUL	SHAKY
LIQUID	FLUID	JUICY	FOREMOST
SACRED	HOLY	TABOO	IDLE
SHINING	GLEAMING	SPARKLING	CONCEALED
WICKED	EVIL	VICIOUS	DAINTY

The stimuli were CVC trigrams of 93–100% Glaze association values (Hilgard, 1951). Two sets of 12 trigrams were selected so as to minimize both interlist and intralist similarity. Only in two instances did syllables on the two lists have two letters in common, and within each list there were 16 repetitions of letters.

Stimuli and responses were combined in four different pairings, making a total of 20 lists. Each list was presented in four different orders. The lists were presented on a Phipps and Bird memory drum at a 2:3 sec. rate, with 6 sec. between trials.

Procedure.—Subjects were seated before the memory drum and read standard paired-associate learning instructions. On the first presentation of each list they were not required to respond, but on subsequent presentations they were encouraged to try to anticipate as many of the words as possible, and it was made clear that there was no penalty for guessing. List 1 was then presented nine times. The list was changed, and *S* was told that he was to learn List 2 "in the same way." The change of lists required 30–40 sec. Nine trials were given on List 2. Following List 2, 30-sec. *Ss* were encouraged to stretch their legs for a few seconds: after approximately 15 sec. instructions for recall were given them. Those *Ss* assigned to the 48-hr. recall groups were told to return in 2 days for the "second hour of the experiment."

The recall instructions differed slightly in

wording for the 30-sec. and 48-hr. groups. For each condition, however, it was emphasized that they were to recall the second list, and that they should begin with the very first syllable they saw. The differences in wording were minor, consisting mainly of greater emphasis on the fact that List 2 was in question for the 48-hr. *Ss*, and presentation of the task as involving possible effects of a "brief pause" for the 30-sec. *Ss*.

Subjects.—The *Ss* were students from the introductory course in psychology at the University of California. They participated in order to fulfill a course requirement, and were naive with respect to verbal learning. They were assigned to conditions at random prior to their appearance, except that 30-sec. and 48-hr. *Ss* were tested in alternation.

Since the experiment was concerned with the effects of List 1, it was felt advisable to discard *Ss* who did not meet a minimal criterion of List 1 learning within the nine trials allowed. On the basis of the pilot *Ss*, it was decided to discard *Ss* who had not achieved at least five correct anticipations on any trial. Seven *Ss* were dropped for not reaching this criterion. An additional 9 *Ss* were dropped for failure to follow instructions; 6 of these admitted not trying to anticipate overtly on early trials, 2 were attempting to learn serial order, and 1 did not anticipate items which had previously been correctly anticipated. While these failures to follow instructions may simply result from the *Ss* in question being slow learners, they could usually do quite well when the instructions were explained to them. An additional 3 *Ss* were dropped for *E* error or apparatus failure. The number of *Ss* discarded for any one of the above reasons in any one similarity condition (pooling 30-sec. and 48-hr. groups) did not exceed 3. A total of 120 *Ss*, 12 in each of the 10 groups, was retained.

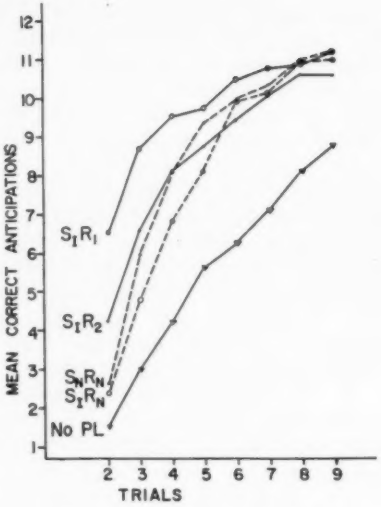

FIG. 1. List 2 learning curves.

TABLE 2

CORRECT RESPONSES AND ERRORS IN LIST 2 LEARNING

Condition		Correct Anticipations		Intra-list Errors %	Inter-list Errors %
		Mean	*SD*		
No PL	30 sec.	47.08	15.71	21.1	—
	48 hr.	41.25	13.24	16.8	—
S_NR_N	30 sec.	69.08	14.80	14.5	0.6
	48 hr.	68.00	8.74	15.6	0.0
S_1R_N	30 sec.	65.42	12.00	7.5	1.4
	48 hr.	62.92	12.51	20.3	0.7
S_1R_2	30 sec.	63.17	15.75	16.1	5.1
	48 hr.	73.75	12.13	14.3	4.6
S_1R_1	30 sec.	77.08	8.28	16.9	18.0
	48 hr.	78.67	9.82	21.0	15.6

RESULTS

List 1 learning.—An analysis of variance of total correct anticipations for the eight anticipation trials of List 1 revealed no differences significant at the .05 level. A similar analysis of intralist errors also revealed no differences. The mean of total correct anticipations for all groups combined was 49.95, with *SD*s ranging from 11.37 to 18.21. For intralist errors, the mean was 8.22, with *SD*s from 2.47 to 7.44. The groups with high response similarity (S_1R_1) made a total of seven interlist errors (words scheduled to appear in List 2) before ever having seen List 2.

List 2 learning.—An analysis of variance of total correct anticipations

similar to that employed for List 1 learning was carried out. Groups scheduled for 30-sec. and 48-hr. recall, not yet differentially treated, did not differ significantly from one another. There were significant differences among the groups as a function of interlist similarity ($F = 23.738$, $df = 4/110$, $P < .005$), and these differences, while small, were in the expected direction (Table 2). All groups showed net positive transfer with respect to the group with no PL. The S_IR_N groups, while showing positive transfer with respect to No PL, were slightly inferior to the S_NR_N groups. This is in agreement with several recent studies which have shown that S_IR_N may produce negative transfer only with respect to a group with equivalent warm-up and learning-to-learn experience (e.g., Besch & Reynolds, 1958; Spiker & Holton, 1958). Examination of the learning curves (Fig. 1) indicates that after the nine trials, all groups with prior learning are within one correct anticipation of one another, while the No PL *S*s reached a criterion comparable to that reached by the other groups on List 1.

Since differences were found among groups in correct anticipations, all error measures were corrected for opportunity. The total number of intralist errors for each *S* was expressed as a proportion of that *S*'s opportunities for error: each stimulus presentation not resulting in a correct anticipation was considered one such opportunity. The resulting proportions were submitted to an arc-sine transformation, and an analysis of variance was carried out. No significant differences in intralist errors were found. In the case of interlist errors, which were extremely infrequent in all groups except those with similar responses (S_IR_1, S_IR_2), the total number of such errors was expressed as a proportion of total opportunities for each group. These data are presented as percentages in Table 2: the percentages, when plotted against decreasing interlist similarity, yield decreasing regular functions, with the 30-sec. groups closely paralleling the 48-hr. groups. A similar gradient results from plotting the number of *S*s making at least one interlist error. In the R_1 groups, 21/24 *S*s made at least one interlist error, while in the R_2 groups 12/24 *S*s made such errors, and in the S_IR_N groups only 5/24 *S*s made interlist errors.

Recall.—In the evaluation of List 2 recall, two methods of estimating differences attributable to differences in List 2 learning have been used. In one, the 30-sec. recall test is used to estimate the strength of List 2 at the end of learning: any difference between 30-sec. and 48-hr. recall scores is attributable to the differential time of testing. A similar estimate results from the other method used, the successive probability analysis (Underwood, 1954, 1956 unpublished[3]). These methods do not allow one to say whether groups which have attained different criteria forget different amounts because of the experimental treatment, or simply because differential amounts of loss are characteristic of different criteria. The ambiguity resulting from different groups having different performance levels at the end of learning arises primarily with respect to comparisons of the group having no PL with the other groups, the No PL *S*s having ended List 2 learning at a lower level of performance. However, there is no pressing reason to suppose that *S*s

[3] Unpublished manuscript by B. J. Underwood entitled, "Strength of association and forgetting."

TABLE 3

CORRECT RESPONSES AND ERRORS IN RECALL

Condition		Correct Anticipations		Intra-list Errors %	Inter-list Errors %
		Mean	*SD*		
No PL	30 sec.	8.66	2.53	16.2	—
	48 hr.	6.17	1.53	9.3	—
S_NR_N	30 sec.	11.00	1.71	8.3	0.0
	48 hr.	7.42	2.27	6.4	0.0
S_IR_N	30 sec.	11.17	1.19	10.0	0.0
	48 hr.	5.33	3.05	2.5	10.0
S_IR_2	30 sec.	9.83	2.04	13.5	9.6
	48 hr.	7.50	2.07	5.5	11.1
S_IR_1	30 sec.	10.92	1.00	19.2	19.2
	48 hr.	7.25	2.09	3.5	18.4

who have reached a lower criterion would ordinarily forget *less* than *S*s who have reached a higher criterion, which appears to have been the case in these data.

Turning first to the comparison of 30-sec. and 48-hr. recall, an analysis of variance revealed that the interaction of retention interval and conditions was significant at the .05 level ($F = 2.822$, $df = 4/110$). The main effects of Conditions ($F = 3.053$, $df = 4/110$) and Retention Interval ($F = 93.04$, $df = 1/110$) were also significant, indicating merely that different conditions performed differently in retention (as in learning), and that forgetting occurred. The mean recall scores are presented in Table 3. They suggest that a good part of the significant interaction may be due to the fact that the loss in S_IR_N is larger than the loss in any other group. When the interaction is partitioned, S_IR_N accounts for most of the variance, the difference among the other groups in amount lost being insignificant.

For the successive probability analysis, the probability of a correct response on Trial 9 following 1, 2, . . . 7 correct anticipations on Trials 2–8 was obtained for each condition, using the combined learning data of the 30-sec. and 48-hr. groups. Since the probability of a correct anticipation following seven correct anticipations was .95 or better in all groups, the probability of a correct response following eight correct anticipations was assigned an arbitrary value of 1.00.

Loss scores were obtained for each 48-hr. *S* individually, by computing an expected recall score on the basis of the probability analysis, and subtracting from this value the obtained recall. These loss scores (Table 4) were subjected to an analysis of variance which confirmed the results of the first analysis of raw recall scores. The conditions differ in amount lost ($F = 4.53$, $df = 4/55$, $P < .01$), and the differential loss appears to result mainly from the greater loss in Cond. S_IR_N, with no difference among the other groups. The logic of the experiment, however, justifies one selected comparison between the No PL control group and the other groups (with S_IR_N excluded). This difference is significant at the .05 level ($F = 4.868$, $df = 1/55$), but since it is a selected comparison, one would probably insist upon significance at the .01 level or better. Finally, a test described by Snedecor (1957, p. 251) which allows for the

TABLE 4

MEANS AND *SD*s OF EXPECTED RECALL, BASED ON PROBABILITY ANALYSIS, MINUS OBTAINED RECALL

Measure	Condition				
	No PL	S_NR_N	S_IR_N	S_IR_2	S_IR_1
Mean	2.61	4.06	6.05	3.88	4.32
SD	1.31	1.97	2.98	1.39	1.97

change in probabilities attendant upon repeated testing of selected groups also indicated that the only significant difference in the loss scores is between S_IR_N and the other groups.

Thus, only in Cond. S_IR_N has PI been demonstrated. This finding is the same as that reported by Morgan and Underwood (1950) with 20-min. recall.

Intralist and interlist errors.—Since errors were not considered to be frequent enough to justify correcting each *S*'s errors, total errors of each kind for each group of *S*s were expressed as a proportion of total opportunities for that group. As in learning, there does not seem to be any obvious systematic relationship between intralist errors and recall.

Interlist errors as a percentage of opportunity *increase* from 30 sec. to 48 hr. for S_IR_N (Table 3) while the two groups with similar responses manifest the same percentage of interlist errors in the two recall tests. The S_NR_N groups made no interlist errors in recall. Since 48-hr. recall scores were lower than 30-sec. recall scores, S_IR_N, S_IR_2, and S_IR_1 all showed an increase in the absolute frequencies of interlist errors, but only in S_IR_N was this increase out of proportion to the increase in opportunities for error.

Transfer and recall as a function of item strength.—In the analyses just presented, the measures employed represented an average of each *S*'s performance on 12 items. It is assumed that such scores indicate what is going on in the learning and recall of individual items. However, few analyses in terms of individual items have been reported (Runquist, 1957). The procedure followed in this experiment, of giving a fixed number of trials on each list, is particularly suited to the analysis of item strength, since an item with a relatively large number of correct anticipations ("reinforcements") is likely to have high strength regardless of whether this strength is the result of its being an easy item or the result of its being learned by a fast learner. This state of affairs should be contrasted with the situation encountered when *S*s learn to criterion. Here, a high number of reinforcements may indicate an easy item or a *slow* learner. Since a slow learner may be assumed to gain less associative strength per reinforcement (Underwood, 1954) there is a confounding of two factors (slow learners and easy items) with opposed effects. Runquist (1957) attempted to reduce this confounding by ranking items within *S*s. Here, no such procedure was felt to be necessary. Not only do the ability of the learner and the ease of the item work in the same way, but an examination of the data revealed that each *S* covered a considerable range of item strengths, with no overwhelming tendency for any one *S* to contribute only weak, or only strong, items.

In preparing Fig. 2 and 3, the number of correct anticipations of an item in List 1 learning was taken as a direct measure of that item's List 1 strength, and the successive probability analysis was used to adjust for List 2 strength. All curves were smoothed by the use of three-point moving averages: the 0 and 8 data are based on averages of two points.

Each of the figures presents recall data as a function of the strength of corresponding items on List 1. A corresponding item, for most groups, was the item with the same stimulus. For Cond. S_NR_N, the correspondence is formal, and reflects (*a*) minor differences in word frequency, minimized for corresponding words, and (*b*) common relative serial positions in the four presentation orders.

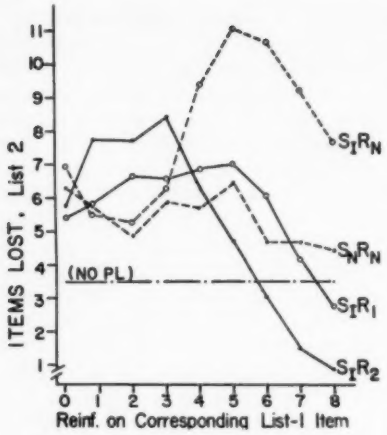

FIG. 2. Recall losses after 48 hr. as a function of List 1 strength.

Figure 2 shows the losses (obtained recall subtracted from the prediction of the probability analysis) in terms of the *number* of items lost. Figure 3 shows percentages of recall, obtained by dividing observed by predicted recall and multiplying by 100. There are no a priori reasons for supposing one way of presenting the data to be better than the other. However, Underwood[3] has shown that in at least one instance, percentage of recall was monotonically related to strength, as expected, while the relationship between loss and strength was nonmonotonic.

Examination of Fig. 2 and 3 reveals that, for some or all degrees of List 1 strength, there is an indication of PI in every group. In Cond. S_IR_N, there appears to be a relative maximum of interference in the vicinity of six reinforcements on List 1. This is in agreement with Underwood's (1945) finding that maximal interference seemed to result when items were of roughly equal strength on the two lists: assuming a normal distribution of item strengths, it would not be unreasonable to expect that most item pairs of roughly equal strength would also be of intermediate strength. The finding is also similar to the familiar fact that, in RI, increased degrees of IL first increase, and then decrease the amount of RI. Particularly interesting is the high degree of forgetting in the groups with similar responses, when the List 1 response is weak. This may help to explain why relatively little PI has been obtained with similar responses: it may be necessary to keep List 1 strength low to obtain maximal interference. There is no reason to expect that the finding of maximal interference with approximate equality of the two lists should be generalizable beyond the S_IR_N case.

DISCUSSION

It will be recalled that Barnes and Underwood (1959) suggested that the unlearning-and-recovery sequence of events applied only to the S_IR_N condition, and that if one assumed no other mechanisms of interference, only S_IR_N would produce PI. However, it was also proposed that with similar responses, loss of differentiation of list membership might lead to forgetting in a delayed test of recall. The data on interlist errors in the present experiment, while not provid-

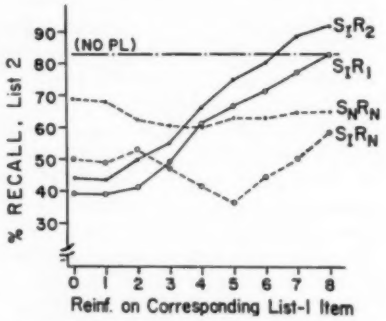

FIG. 3. Percentage recalled as a function of List 1 strength.

ing a direct test of either mechanism, have an important bearing on this question. All groups with identical stimuli show an increase in interlist errors from 30 sec. to 48 hr. Such errors are rare in Cond. S_NR_N, and are not made at all in recall. An increase in the *number* of interlist errors may or may not reflect loss of differentiation, depending on whether one grants primary importance to the error, or to the failure to respond correctly which made an error possible. If one assumes that *S* first loses the correct response, and then may or may not make an error, then (using the correction for opportunity), the data indicate increased interlist generalization only for S_IR_N, the other groups continuing to contribute interlist errors in the same proportion to opportunities at 48 hr. as at 30 sec. If, on the other hand, one assumes that the error displaces a correct response which would otherwise be given, then the appropriate comparison is between the absolute number of errors, which increase in all groups. This would leave unexplained the fact that with similar responses, the increase is proportional to the forgetting obtained, while in S_IR_N it is not. The author's preference is to assume that loss of differentiation is of relatively minor importance, as suggested by the overall equality of total forgetting in all groups except S_IR_N, and that the disproportionate increase of errors in the S_IR_N condition is an indication of recovery of List 1 responses. Such a conclusion is in agreement with the implications of Barnes and Underwood's data, and should suffice as a conservative explanation of the present experiment.

Two things remain unexplained: the indication of some PI in the S_NR_N item analysis, and the apparent interaction of PI and List 1 strength in the similar-response conditions. In the case of S_NR_N, it is clear that the total stimulus contexts of the two lists are not completely dissimilar—in fact, letters common to the two sets of stimuli may have been sufficient to generate a marginal amount of interlist interference. Other common stimuli involve the room, *E*, the memory drum, etc. . .

In the similar-response groups, the mediation hypothesis leads one to suspect that the items readily forgotten were those with weak mediators, while the items well recalled were those with strong mediators. Weak mediators might be eliminated by extraexperimental interference, and may provide ideal conditions for the confusion of the two responses. Thus, while loss of differentiation seems not to explain the data as a whole, it may interact with the mediation mechanism, being effective only with low List 1 strength. The explanation of List 2 forgetting as a result of the loss of the List 1 mediator runs into several difficulties. First, the transfer data suggest that when a List 1 item is weak, the dependence of the List 2 response upon it should be minimal—and the loss of an ineffective mediator would be of little consequence. Furthermore, the mediation involved may not be a simple response-chaining process in which the List 1 response serves as a cue for the List 2 response. If a weak List 1 response is strengthened during List 2 learning (as RI experiments seem to suggest), then an interdependence of List 1 and List 2 responses could be built up which would be more complex than the simple response chain assumed to occur when a strong mediator is available. It is obvious that further study of the similar response conditions is necessary before any conclusions can be reached regarding these possibilities.

The data of the present experiment once more indicate the usefulness of considering the unlearning-and-recovery sequence as a basic process in RI and PI, and suggest that while interlist similarity per se is not a highly potent variable, it may yet be found to produce large effects in interaction with degree of learning.

Summary

Failures to demonstrate an effect of interlist response similarity on PI have previously been reported in experiments in which short retention intervals have been used, with relatively small amounts of PI. In the experiment reported, an attempt was made to

maximize the chances of obtaining PI by using a 48-hr. retention interval.

The design involved 10 groups of *Ss*, each of which learned a common List 2 after learning first lists which differed in their similarity to List 2. The first lists had identical stimuli and similar, less similar, or dissimilar responses in three of the basic conditions. Two further conditions involved groups which learned either no List 1, or a List 1 in which both stimuli and responses were dissimilar to those on List 2. In each condition, retention of List 2 was tested after 30 sec. and 48 hr., with independent groups tested at each time interval.

The results showed significant PI only for the condition in which the two lists had identical stimuli and dissimilar responses. However the degree of List 1 strength associated with maximal interference was different for each condition, suggesting that significant PI might be obtained in each condition by appropriate manipulations of List 1 strength.

References

Barnes, J. M., & Underwood, B. J. "Fate" of first-list associations in transfer theory. *J. exp. Psychol.*, 1959, **58,** 97–105.

Besch, N. F., & Reynolds, W. F. Associative interference in verbal paired-associate learning. *J. exp. Psychol.*, 1958, **55,** 554–558.

Haagen, C. H. Synonymity, vividness, familiarity, and association-value ratings of 400 pairs of common adjectives. *J. Psychol.*, 1949, **30,** 185–200.

Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.

Morgan, R. L., & Underwood, B. J. Proactive inhibition as a function of response similarity. *J. exp. Psychol.*, 1950, **40,** 592–603.

Postman, L. The present status of interference theory. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior*. New York: McGraw-Hill, 1961.

Runquist, W. N. Retention of verbal associates as a function of strength. *J. exp. Psychol.*, 1957, **54,** 369–374.

Slamecka, N. J., & Ceraso, J. Retroactive and proactive inhibition of verbal learning. *Psychol. Bull.*, 1960, **57,** 449–475.

Snedecor, G. W. *Statistical methods*. Ames: Iowa State Coll. Press, 1957.

Spiker, C. C., & Holton, R. B. Associative transfer in motor paired-associate learning as a function of amount of first-task practice. *J. exp. Psychol.*, 1958, **56,** 123–132.

Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words*. New York: Teacher's College, Columbia University, 1944.

Underwood, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychol. Monogr.*, 1945, **59**(3, Whole No. 273).

Underwood, B. J. Speed of learning and amount retained: A consideration of methodology. *Psychol. Bull.*, 1954, **51,** 276–282.

Young, R. K. Retroactive and proactive effects under varying conditions of response similarity. *J. exp. Psychol.*, 1955, **50,** 113–119.

(Received August 29, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 373–379

EFFECTS OF NONREINFORCED TRIALS IN TWO-CHOICE LEARNING WITH NONCONTINGENT REINFORCEMENT [1]

JAMES G. GREENO [2]

University of Minnesota

The purpose of this paper is to examine the role of nonreinforced trials in a simple prediction situation. Stimulus sampling theories have successfully accounted for many of the results from choice experiments in which some one of the alternative responses is reinforced on each trial (see Estes, 1959). However, these theories have not yet been extended so as to provide an adequate account of results from experiments which include trials on which no reinforcement is presented.

In the present studies, *S* was instructed to predict which of two lights would flash on each of a series of trials. Following *S*'s choice on each trial, one of three events occurred. On some trials, one of the lights flashed (E_1 or E_2) constituting a reinforcement for the response of predicting the light that flashed (A_1 or A_2). On other trials, neither light flashed (E_0) constituting a nonreinforced trial. The proportions of trials on which E_1, E_2, and E_0 occurred are denoted π_1, π_2, and π_0, respectively. In this paper, p_1 will designate the observed proportion of A_1 choices.

Two hypotheses have been offered frequently to account for the effects of E_0 trials. One of these is the identity hypothesis, which asserts that an E_0 trial leaves choice probabilities unchanged. A second suggestion is the correction hypothesis, which implies that an E_0 trial reduces the probability of the response chosen by *S* on that trial. Estes (1959) summarized the results of several experiments involving E_0 trials and suggested that the correction and identity hypotheses describe processes which occur in different degrees in different situations, depending upon such variables as instructions and preliminary training.

In particular, results obtained by Atkinson (1956), Neimark (1953), and Millward (1960) are consistent with identity-hypothesis predictions. However, Anderson and Grant (1957, 1958) and LaBerge, Greeno, and Peterson (1962) have obtained results indicating that with $\pi_1 > \pi_2$, E_0 trials reduce p_1; and sequential statistics reported in these studies suggest that the obtained changes in p_1 may not have been produced by a correction effect.

In most studies of the effects of E_0 trials, investigators have tested for quantitative agreement between data and predictions from specific theories. On the other hand, the present study is intended to provide evidence regarding the qualitative properties of the effect of E_0 trials. Therefore, E_0 trials were presented in situations for which ordinal predictions could be derived so as to differentiate between

[1] These results were included in the author's dissertation, presented to the faculty of the Graduate School, University of Minnesota, in partial fulfillment of the requirements for the PhD. Thanks are due to D. L. LaBerge, who served as major advisor for the dissertation and provided a critical reading of this paper. Marianne Larson and Muriel Dieteman assisted with data analysis. While conducting these studies, the author received financial support from the Ford Foundation Behavioral Science Training Program and from the Danforth Foundation.

[2] Now at Indiana University.

hypotheses. In terms of empirical variables, the present experiments were not designed primarily to test the quantitative effects of experimental operations. Rather, these data were obtained in order to determine whether certain variables are relevant in relation to the effects of E_0 trials.

Experiment I

This study was designed to investigate further the finding that E_0 trials reduce p_1. Evidence was sought regarding two questions:

1. Is the effect of an E_0 trial invariant with respect to the number of trials on which *S* has received E_0? This question arises from the possibility that the effect of E_0 trials might decrease over a series of partially reinforced trials. This might occur if E_0 events acquired secondary reinforcing properties through association with reinforced trials in the sequence (Bush, 1960), or if the effect of E_0 trials depended upon disrupting *S*'s behavior and disruption effects diminished as E_0 events continued to occur (Neimark, 1953).

2. Do E_0 trials reduce the asymptotic value of p_1? LaBerge, Greeno, and Peterson's (1962) results included differences among mean choice frequencies due to E_0 trials during acquisition. This second question then, simply asks whether such a difference also occurs during near-asymptotic performance.

Method

The 72 *Ss* were students in introductory psychology classes. Two *Ss* at a time were seated in a darkened room and instructed to predict which of two red lights would flash each time they heard a buzzer signal. Each *S* indicated his prediction on each trial by pressing one of two spring-release levers. The *Ss* received eight practice trials: $E_1E_1E_2E_2E_0E_0E_1E_2$. The *E* emphasized that *S* "should make a choice on each trial, no matter what happens," and then asked for questions. If *S* asked or remarked about the E_0 events, *E* said, "That may happen on some trials. In any case you should make a choice on each trial." No other instructions were given regarding E_0 trials. Following the instructions, trials were presented without interruption until all trials had been presented.

Each experimental trial consisted of the following events: buzzer signal, 2 sec.; off, 1 sec.; reinforcement event, 1 sec.; off, 2 sec. This temporal sequence was automatically controlled. On each trial, *S*'s choice and the reinforcement event were automatically recorded.

Reinforcement schedules were constructed by randomly ordering 20-trial blocks of E_1 and E_2 trials, and then randomly adding the number of E_0 trials needed to satisfy the specified value of π_0. Each block, then, included 20 reinforced trials, and a total of $20/(1 - \pi_0)$ trials.

The three experimental groups received reinforcement sequences as follows:

Group 0—
Blocks 1–10: $\pi_1:\pi_2 = 90:10$; $\pi_0 = 0$.

Group 50—
Blocks 1–10: $\pi_1:\pi_2 = 90:10$; $\pi_0 = .50$.

Group 50*P*—
Blocks P_1–P_2: $\pi_1:\pi_2 = 50:50$; $\pi_0 = .67$,
Blocks 1–5: $\pi_1:\pi_2 = 90:10$; $\pi_0 = .50$.

A comparison between Groups 50 and 50*P* is relevant to the question of invariance of the effects of E_0 trials. The pretraining schedule received by Group 50*P* was selected because LaBerge, Greeno, and Peterson (1962) found that this sequence produced no change in p_1 for a 90:10 sequence with $\pi_0 = 0$. Thus, if Groups 50 and 50*P* were to differ over Blocks 1–5 of the present study, then this difference could be attributed to a change in the reinforcing effect of E_0 trials due to their presence in the pretraining sequence.

A comparison of Groups 0 and 50 over Blocks 6–10 is relevant to the question regarding near-asymptotic properties of E_0 effects.

Six different random schedules were used, with orders of E_1 and E_2 events matched across groups for comparable blocks. The right- and left-hand response, respectively, was designated A_1 for one-half of the *Ss* in each condition.

Results and Discussion

The mean proportions of A_1 choices (p_1) for each group are graphed by

blocks in Fig. 1. Over the first five blocks of 90:10 reinforcement, p_1 was obtained for each S. An estimate of $\sigma^2 = .0106$ ($df = 36$) was obtained as the residual mean square of the $3 \times 2 \times 6$ factorial analysis of variance involving experimental groups, right or left sides, and schedules as the factors. Then 90% confidence intervals for orthogonal contrasts between group means were estimated as follows:

$$\bar{p}_1(0) - \frac{\bar{p}_1(50) + \bar{p}_1(50P)}{2} = .091 \pm .052;$$

$$\bar{p}_1(50P) - \bar{p}_1(50) = -.018 \pm .060.$$

In qualitative terms, p_1 was less with $\pi_0 = .50$ than with $\pi_0 = 0$; and p_1 with $\pi_0 = .50$ did not differ significantly as a result of the 50:50 pretraining trials. The difference due to E_0 trials over these blocks replicates earlier findings and thus provides additional contraindication for the identity hypothesis. Since there was not a significant increase in p_1 due to the 50:50 pretraining, there is no evidence that the effect of E_0 trials decreased over the partially reinforced sequence of 50:50 trials.

In order to obtain evidence as to whether E_0 trials influenced asymptotic values of p_1, the value of p_1 was obtained for each S in Groups 0 and 50 for each block during Blocks 6–10. These scores were subjected to analysis of trends (Grant, 1956). The F for overall linear trend (1.36; $df = 4/184$; $P > .25$) and the F for difference between group linear trends (1.16; $df = 1/46$; $P > .75$) were not significant. Therefore, there is no statistical evidence that Groups 0 and 50 were not at asymptote during Blocks 6–10. The F for difference between group means was significant (9.28; $df = 1/46$; $P < .0005$), indicating that the asymptotic value of p_1 was changed by the E_0 trials received by Group 50.

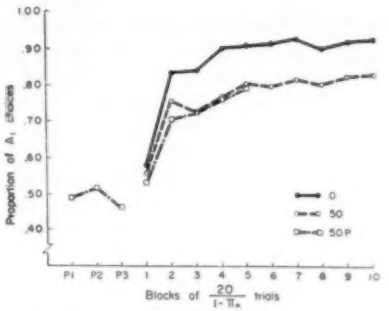

FIG. 1. Mean proportions of A_1 choices across blocks of trials in Exp. I. (In Blocks *P1–P3*, $\pi_1:\pi_2 = 50:50$. In Blocks 1–10, $\pi_1:\pi_2 = 90:10$.)

In Table 1 are presented estimated means and variances of p_1 scores, and proportions of events ($p_{j,ik}$) such that A_i and E_j occurred on Trial n, and A_k occurred on Trial $n + 1$. Lines 1 and 2 of Table 1 are relevant to the present discussion. Anderson and Grant (1958) used a statistic based on similar data to estimate changes in p_1 across E_0 trials. Let $\gamma = p_{0,21}/\pi_0 - p_{0,12}/\pi_0$. For Group 50 of the present study, $\gamma = -.034$, indicating that p_1 decreased across E_0 trials. This result is consistent with those obtained in Anderson and Grant's analysis, and suggests that the change in $\bar{p}_1$ reported here was due to effects of E_0 trials on choice probabilities, rather than to changes in the effects of E_1 and E_2 events.

Taken together, the obtained difference in $\bar{p}_1$ and the negative value of γ provide evidence against the identity hypothesis which is particularly compelling, since it is based on near-asymptotic performance.

EXPERIMENT II

This experiment was designed to provide evidence relevant to the cor-

rection hypothesis, which implies that the effect of an E_0 trial depends upon *S*'s choice on the trial. Therefore, if the correction hypothesis were correct, then the average effect of an E_0 trial would depend upon response frequencies. An alternative possibility is that the average effect of an E_0 trial might depend upon reinforcement frequencies.

In the standard prediction situation with noncontingent reinforcement, it is impossible to discriminate between the effect of these variables on the basis of ordinal hypotheses, since response and reinforcement frequencies turn out to be equal at asymptote. In order to remove this equality, unequal incentives were introduced for the two response alternatives. Under these conditions, *S* should choose the response associated with the higher incentive more frequently than that response is reinforced. Thus, it was hoped that the contributions of response frequencies and reinforcement frequencies to the effect of E_0 trials could be separated.

Optimal conditions for this purpose would include a condition for which $p_1 = .50$ when $\pi_1 > \pi_2$ and $\pi_0 = 0$, and another condition for which $p_1 < .50$ with $\pi_1 = \pi_2$ and $\pi_0 = 0$. The simplest form of the correction hypothesis would predict that E_0 trials should increase p_1 in the second condition, and that E_0 trials would have no effect in the first condition. On the other hand, if the effect of E_0 trials depended upon reinforcement frequencies, there should be no effect due to E_0 trials in the second condition, and p_1 should be decreased by E_0 trials in the first condition.

The incentive operation used was to instruct *S* that he would receive more points for a correct prediction of one light than the other. Preliminary studies indicated that the optimal conditions described above would be best approximated with a point ratio of 1:6.

Method

The 250 *S*s were students in introductory psychology classes. Procedures and apparatus were the same as those used in Exp. I with the following exceptions: Instead of facing a box with red lights and levers, *S* faced a slanting panel which held two spring-release buttons with which he indicated his choices. The signal light above the left button was red, and the light above the right button was white. The numerals 1 and 6 were displayed beside the left and right buttons, respectively.

Instructions indicated that the experiment was a test of *S*'s skill at making choices. The task for *S* was to get as many points as he could by predicting which light would flash on each trial; and *S* was told that he would receive one point each time he correctly predicted the red left-hand light and six points each time he correctly predicted the white right-hand light. Practice trials and remarks about E_0 trials were as in Exp. I.

This experiment consisted of three sections which were run separately. In Section 75, $\pi_1:\pi_2 = 75:25$; in Section 50, $\pi_1:\pi_2 = 50:50$; in Section 25, $\pi_1:\pi_2 = 25:75$. In each section, a group with $\pi_0 = 0$ (Groups 75/0, 50/0, 25/0) was compared with a group with $\pi_0 = .50$ (Groups 75/50, 50/50, 25/50).

In all cases, six blocks of trials were presented. Eight different random schedules were used. Schedules were constructed in the manner described for Exp. I. Schedules were matched within sections with respect to the order of E_1 and E_2 events. Schedules were matched across sections with respect to the trial numbers on which E_0 occurred for the groups with $\pi_0 = .50$. The left-hand response was designated A_1 for all *S*s.

For Sections 75 and 50, a double sampling technique was used to determine the number of *S*s to be used (Cox, 1958). In each case, n was set so that the difference between mean values of p_1 for π_0 conditions could be estimated with a 90% confidence interval of length less than .10. The number of *S*s in the six experimental groups were as follows: Group 75/0, 64; Group 75/50, 32; Groups 50/0 and 50/50, 48; Groups 25/0 and 25/50, 32.

Results and Discussion

Mean values of p_1 are graphed by blocks in Fig. 2. The value of p_1 over

Blocks 3–6 was obtained for each S and these scores were analyzed. Estimates of σ^2 were obtained as residual mean square terms of factorial analyses of variance (see Table 1). Using these estimates, 90% confidence intervals were estimated for the differences between $\pi_0 = 0$ and $\pi_0 = .50$ conditions as follows:

$$\tilde{p}_1(75/0) - \tilde{p}_1(75/50) = .079 \pm .046;$$
$$\tilde{p}_1(50/0) - \tilde{p}_1(50/50) = .053 \pm .048;$$
$$\tilde{p}_1(25/0) - \tilde{p}_1(25/50) = -.100 \pm .050.$$

Lines 3–6 of Table 1 provide more detailed information regarding performance of these groups. The values of the γ statistic described above were estimated as follows: Group 75/50, $-.039$; Group 50/50, $+.006$; Group 25/50, $+.030$. These estimates suggest the following: p_1 decreased across E_0 trials with $\pi_1:\pi_2 = 75:25$; p_1 increased across E_0 trials with $\pi_1:\pi_2 = 25:75$; and p_1 was virtually unchanged by E_0 trials with $\pi_1:\pi_2 = 50:50$.

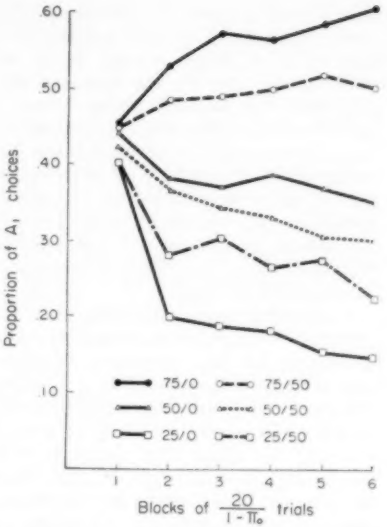

FIG. 2. Proportion of A_1 choices across trial blocks in Exp. II.

Two formulations of the correction hypothesis could be applied to these data, although neither of them seems

TABLE 1

ESTIMATES OF GROUP MEANS AND VARIANCES AND PROPORTIONS OF RESPONSE-REINFORCEMENT COMBINATIONS

Group	$\tilde{p}_1$	$\sigma^2(p_1)$	$p_{1,11}$	$p_{1,12}$	$p_{2,11}$	$p_{2,12}$	$p_{0,11}$	$p_{0,12}$	$p_{1,21}$	$p_{1,22}$	$p_{2,21}$	$p_{2,22}$	$p_{0,21}$	$p_{0,22}$
	Experiment I, Blocks 6–10													
0	.917	.0029	.772	.052	.084	.008	—	—	.052	.022	.008	.001	—	—
50	.814		.330	.038	.035	.007	.334	.068	.059	.023	.006	.003	.051	.046
	Experiment II, Blocks 3–6													
75/0	.584	.0242	.310	.135	.085	.056	—	—	.141	.165	.046	.063	—	—
75/50	.505	.0129	.138	.056	.030	.030	.126	.130	.079	.099	.027	.037	.111	.136
50/0	.374	.0199	.093	.095	.082	.103	—	—	.112	.201	.094	.221	—	—
50/50	.321		.047	.037	.036	.047	.082	.083	.043	.136	.039	.128	.086	.246
25/0	.169	.0139	.015	.029	.024	.103	—	—	.045	.160	.082	.543	—	—
25/50	.269		.013	.018	.028	.076	.051	.085	.026	.070	.052	.220	.100	.261

Note.—See text for explanations of entries.

adequate to account for the results that were obtained. First, we could expect that E_0 trials might change p_1 in the direction of .50 by an amount proportional to the difference between $p_1(\pi_0 = 0)$ and .50. If this had been the case, however, then the difference between $\bar{p}_1(75/0)$ and $\bar{p}_1(75/50)$ would seem to have been too large. Instead of decreasing p_1 toward .50, the E_0 trials reduced p_1 *to* .50. More critically, in Section 50, $\bar{p}_1(\pi_0 = .50)$ was farther from .50 than $\bar{p}_1(\pi_0 = 0)$.

A second formulation would allow the probability of a correction response to be influenced by the incentive variable. Such a formulation would be consistent with results indicating that the incentive variable had a greater effect in groups receiving E_0 trials than with $\pi_0 = 0$. The results from Sections 75 and 50 of the present experiment are, therefore, consistent with such a formulation. However, had the incentive operation been more effective in Group 25/50 than Group 25/0, there should not have been a significant difference between these groups in the obtained direction.

In Sections 75 and 50, then, E_0 trials reduced the frequency with which *S*s chose that response which was reinforced more frequently, although the more frequently reinforced response was associated with different incentives in the two cases. In Section 50, although $\bar{p}_1$ was changed by the presence of E_0 trials, it appears that p_1 did not change across E_0 trials despite the fact that $p_1 < .50$. These results, then, indicate that *S*'s choice frequency is not a relevant variable in relation to the effect of E_0 trials; and the present data therefore contraindicate the correction hypothesis. On the other hand, the present findings indicate that the effect of E_0 trials is related to the relative frequencies with which the alternative choices are reinforced.

Summary

Two experiments were conducted in order to obtain evidence regarding the effect of nonreinforced trials in a simple prediction situation. Evidence was obtained regarding two hypotheses: (*a*) The identity hypothesis, which asserts that an E_0 trial does not change choice probabilities; and (*b*) The correction hypothesis, which implies that an E_0 trial reduces the probability of the response chosen by *S* on that trial.

In Exp. I, two groups receiving sequences of trials including 200 reinforcements with $\pi_1:\pi_2 = 90:10$ were compared over trial blocks during which p_1 was near its asymptote. The group receiving E_0 trials showed a lower value of p_1 than did the group with $\pi_0 = 0$. This result contraindicates the identity hypothesis. A second comparison from Exp. I examined the effect of a pretraining sequence including E_0 trials with $\pi_1:\pi_2 = 50:50$. The pretraining sequence did not significantly change the effect of E_0 trials in the 90:10 sequence which followed. There was, then, no evidence that the effect of E_0 trials decreases over sequences of partially reinforced trials.

In Exp. II, unequal incentives were offered *S* for the two choice responses in an attempt to separate the contributions of choice frequencies and reinforcement frequencies to the effect of E_0 trials. It was found that with $\pi_1:\pi_2 = 75:25$, E_0 trials reduced p_1; with $\pi_1:\pi_2 = 25:75$, E_0 trials increased p_1. With $\pi_1:\pi_2 = 50:50$, $p_1(\pi_0 = .50)$ was less than $\bar{p}_1(\pi_0 = 0)$, although a sequential statistic indicated that p_1 did not change across E_0 trials. In each case, $p_1(\pi_0 = 0)$ was less than $\frac{\pi_1}{\pi_1 + \pi_2}$. Therefore, it was concluded that the effect of E_0 trials was related to the frequencies with which the choices were reinforced, rather than to the frequencies with which *S*s chose the responses. Since the correction hypothesis implies that choice frequencies determine the average effect of E_0 trials, the results of Exp. II were interpreted as a disconfirmation of the correction hypothesis.

References

Anderson, N. H., & Grant, D. A. A test of a statistical learning theory model for two-choice behavior with double stimulus events. *J. exp. Psychol.*, 1957, **54**, 305–317.

Anderson, N. H., & Grant, D. A. Correction and reanalysis. *J. exp. Psychol.*, 1958, **56**, 453–454.

Atkinson, R. D. An analysis of the effect of nonreinforced trials in terms of statistical learning theory. *J. exp. Psychol.*, 1956, **52**, 28–32.

Bush, R. R. A survey of mathematical learning theory. In R. D. Luce (Ed.),

Developments in mathematical psychology. Glencoe, Ill.: Free Press, 1960. Pp. 120–165.

COX, D. R. *Planning of experiments.* New York: Wiley, 1958.

ESTES, W. K. The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: A study of a science.* Vol. 2. *General systematic formulations, learning, and special processes.* New York: McGraw-Hill, 1959. Pp. 380–491.

GRANT, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.

LABERGE, D., GREENO, J. G., & PETERSON, O. F. Nonreinforcement and neutralization of stimuli. *J. exp. Psychol.*, 1962, **63**, 207–213.

MILLWARD, R. B. A comparison of two learning models for two-choice conditioning experiments involving nonreinforced trials. Unpublished doctoral dissertation, Indiana University, 1960.

NEIMARK, E. D. Effects of type of nonreinforcement and number of alternative responses in two verbal conditioning situations. Unpublished doctoral dissertation, Indiana University, 1953.

(Received September 1, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 380–387

RETENTION OF FIRST-LIST ASSOCIATIONS AS A FUNCTION OF THE CONDITIONS OF TRANSFER[1]

LEO POSTMAN

University of California

A basic question for theories of transfer and retroactive inhibition (RI) is whether the strength of first-list associations changes systematically during the learning of the second list. A recent study by Barnes and Underwood (1959) has presented clear evidence that the "fate" of first-list associations depends on the conditions of intertask transfer. When the paradigm for negative transfer (identical stimuli and dissimilar responses—A-B, A-C) is used, List 1 associations are unlearned or extinguished during the acquisition of List 2. When the successive tasks conform to the paradigm for positive transfer (identical stimuli and highly similar responses—A-B, A-B′), List 1 associations are maintained at high strength and appear to mediate the reproduction of List 2 responses. The present paper presents additional findings in support of these conclusions.

Experimental evidence consistent with the unlearning hypothesis has been accumulating steadily but has remained short of decisive prior to the study of Barnes and Underwood (Postman, 1961). The hypothesis was first advanced by Melton and Irwin (1940) as part of a two-factor theory of RI. According to this theory, List 1 associations which have been unlearned are not available at the time of recall; others which remain potentially available are displaced by competing associations from List 2. The lack of correlation between total amount of RI and the number of overt interlist intrusions is in accord with this interpretation as is the finding that RI is greater than proactive inhibition (PI) at short retention intervals (Melton & Von Lackum, 1941; Underwood, 1948a). The assumption that some List 1 associations are not available to *S* at the time of the retention test received further support from studies using the method of modified free recall or MFR (Briggs, 1954; Briggs, Thompson, & Brogden, 1954; Underwood, 1948b). After learning two successive lists (A-B, A-C), *S* is presented with the common stimulus term (A) and is required to give either the List 1 response (B) or the List 2 response (C). The relative frequency of List 1 responses declines steadily as a function of the degree of List 2 learning. The fact that the proportion of List 1 responses increases as a function of time supports the interpretation of unlearning as a process akin to extinction followed by spontaneous recovery. The results of MFR tests do not, however, provide crucial evidence for reduced availability of List 1 associations. Since *S* is instructed to give either B or C, a progressive rise in the proportion of C responses may merely signal increasing dominance of List 2 over List 1 associations and does not compel the conclusion that the latter are not available to *S*.

A critical test of the unlearning hypothesis requires that the availability of List 1 associations be assessed under conditions in which the effects of response competition and of losses in list differentiation are eliminated. This requirement was met for the first time in the study of Barnes and Underwood (1959). In the experiment using the A-B, A-C paradigm with nonsense syllables as stimuli and adjectives as responses, *S*s were presented with the common stimulus terms at the end of List 2 learning and were required to write down both the List 1 and List 2 responses to each of the stimuli. Following a usage adopted by Melton (1961) this modification of MFR will be referred to as MMFR. With List 1 learned to a criterion of one perfect recitation, the MMFR test was administered to different groups after 1, 5, 10, or 20 anticipation trials on List 2. As degree of List 2 learning increased, there was a steady decline in the number of List 1 responses and an equally regular rise in the number of List 2 responses. These trends were obtained regardless of whether credit was given for all responses recalled or only those which were reproduced to the appropriate stimulus and identified

[1] This research was supported by a grant from the National Science Foundation.

correctly as to list membership. With the exception of the lowest degree of List 2 learning, List 1 responses were reproduced before List 2 responses with only chance or less than chance frequency. These findings give unequivocal support to the unlearning hypothesis and are inconsistent with the "independence hypothesis," according to which the two systems of responses remain independent and intact and RI is due to reproductive inhibition.

The inverse relationship between interlist response similarity and RI suggests that List 1 responses gain in strength during the acquisition of List 2 when intertask transfer is positive. One theoretical account of this relationship assumes that there is generalization of reinforcement between List 1 and List 2 (Osgood, 1946, 1948; Underwood, 1951; Young, 1955). While this hypothesis accounts for the relationship between positive transfer and RI, a second experiment by Barnes and Underwood supports a different conception of the "fate" of List 1 associations under the A-B, A-B′ paradigm. In a replication of the procedure described above with lists conforming to the A-B, A-B′ paradigm, the MMFR test showed little decline in the recall of List 1 associations as a function of the degree of List 2 learning. Recall of List 2 associations was nearly perfect after only one trial. At the lower degrees of List 2 learning, List 1 responses tended to be recalled before List 2 responses. The total pattern of results, and especially the almost instantaneous acquisition of List 2, points to direct mediation of List 2 responses by List 1 responses (A-B-B′) rather than generalization of reinforcement as the mechanism responsible for the positive transfer effects. Since similarity and associative connection between items are correlated, rehearsal of A-B-B′ strengthens List 1 responses and at the same time leads to a high level of recall for List 2 responses. Most *S*s in the experiment of Barnes and Underwood reported using mediation during List 2 learning.

The study of Barnes and Underwood (1959) represents an important advance in the analysis of the mechanisms of transfer which are fundamental to an interference theory of forgetting. The conclusions concerning the "fate" of List 1 associations will be strengthened if it is possible to show that the differences obtained with the two paradigms are not a function of *S*'s set at the time of recall. Results obtained by the method of anticipation suggest that RI at recall may be enhanced by *S*'s tendency to continue responding from the list which he practiced last. Such "generalized competition" leads to a large proportion of failures to respond on the test of recall even when the two lists do not share common stimulus terms and List 2 has not been learned to a higher degree than List 1 (Newton & Wickens, 1956; Postman, 1961; Postman & Riley, 1959). It is reasonable to suppose that generalized competition is greater when intertask transfer is negative than when it is positive, especially if there is direct mediation of responses under the latter condition. When recall for homogeneous lists is tested by MMFR, differential effects of set on the reproduction of List 1 responses cannot be ruled out. A mixed-list design was used in the present study in order to assess the differences in the availability of List 1 responses for the A-B, A-C and A-B, A-B′ paradigms with the effects of response set equalized. The need for control of response set in MMFR by means of mixed-list designs has been pointed out by Melton (1961).

Performance in MMFR is relatively free from the effects of response competition and list differentiation. A comparison of the amounts of RI as measured by MMFR and the conventional anticipation method will permit an estimate of the extent to which conventional measures of RI reflect reduction in the availability of List 1 responses on the one hand, and response competition and loss of list differentiation on the other. A conventional test of anticipation and MMFR were, therefore, used with different groups to measure RI under the conditions of the present experiment.

Method

Experimental design.—Four groups of *S*s—two Work groups and two Rest groups—learned a list of eight paired associates (A-B) to a criterion of one perfect recitation. The Work groups were then given 20 trials on a second list of eight paired associates. For half of the pairs in List 2 the relationship to the pairs in List 1 conformed to the A-B, A-C paradigm, and for the other half of the pairs to the A-B, A-B′ paradigm. The Rest groups rated a series of pictures on several evaluative dimensions for a period equal to that spent in List 2 learning by the Work groups. At the end of the retention interval one Work group

and one Rest group were given an MMFR test in which the common stimulus terms (A) were presented and *Ss* were required to give both List 1 and List 2 responses. Following the MMFR test List 1 was relearned for 10 trials. The other Work group and Rest group relearned List 1 for 10 trials without an intervening MMFR test. The four conditions included in the design will be designated as MMFR Work, MMFR Rest, Conventional Work, and Conventional Rest.

Lists.—The stimulus terms were eight nonsense syllables of 47–53% association value (Glaze, 1928). The intrastimulus similarity was low. None of the consonants were duplicated, and each of four vowels was repeated once. The pool of response terms consisted of two-syllable adjectives from Haagen's (1949) tables. There were eight sets of three adjectives each. Two of the three adjectives in each set were those used by Barnes and Underwood (1959) as responses in A-B, A-B′ pairs and had similarity ratings from .9 to 1.4 on Haagen's scale. The third adjective in each set had no apparent relationship to the other two. Pairs were assigned to lists so that (*a*) each nonsense syllable was the common stimulus for similar responses half the time and for dissimilar responses the other half of the time, and (*b*) similar responses and dissimilar responses from a given set of adjectives were each paired with a common stimulus half the time. There were two combinations of lists; the two lists within each combination were learned first and second equally often. Intralist response similarity was low throughout: there were no duplications of first letters, and no more than one duplication of a terminal suffix.

Procedure.—The lists were presented on a Hull-type memory drum at a 2:2 rate, with an 8-sec. intertrial interval. There were four different orders of pairs each of which was used as a starting order equally often. The two lists learned by the Work groups were separated by 2 min. The total retention interval, which was filled by the picture-rating task for the Rest groups, was 15.3 min.

For the MMFR test the drum was operated manually and the exposure of successive stimulus terms was paced by *S*. The *Ss* were instructed to call out the two responses in the order in which they occurred to them when the common stimulus term appeared in the window.

Subjects.—With two conditions (Work vs. Rest) and two tests of recall (MMFR vs. Conventional), there were four groups of 16 *Ss* each. The *Ss* were undergraduate students who were not necessarily naive to rote-learning experiments but had no previous experience with MMFR tests. For purposes of assignment to conditions, 64 entries were made so that to each *S* in a Work group there corresponded an *S* in the appropriate Rest group who learned the same test list with the same starting order. The *Ss* were run in blocks of 4, with 1 *S* per block drawn at random from each of the four conditions. The running order within blocks was determined by a table of random numbers. No *Ss* were lost because of failure to learn.

Results

List 1 learning.—The mean number of trials to criterion on List 1 for the combined groups was 13.26, with an *SD* of 7.98. The means for individual groups ranged between 12.81 and 14.00 and did not differ significantly ($F < 1$).

The mean number of trials to a criterion of 4/4 for the pairs used in the A-B, A-B′ paradigm was 10.39, and 10.31 for the pairs used in the A-B, A-C paradigm. For purposes of this comparison as well as in subsequent analyses of recall and relearning, the classification of the pairs learned by *Ss* in the Work groups was applied to the protocols of corresponding *Ss* in the Rest groups.

List 2 learning.—During practice on List 2, the A-B′ paradigm produced faster learning than the A-C paradigm. The mean number of correct responses on A-B′ pairs was 65.06 ($SD = 12.88$), and 59.47 ($SD = 11.15$) on A-C pairs. The results for the Work groups were quite similar—64.94 vs. 59.31, and 65.19 vs. 59.62 for the MMFR and Conventional groups, respectively. The difference between the two kinds of pairs is significant ($t = 2.45$, $df = 31$, $.01 < P < .02$) and is in accord with the differential transfer effects normally predicted for the two paradigms.

The number of interlist intrusions given to the appropriate stimulus terms was greater for the A-B, A-B′

TABLE 1

NUMBER OF LIST 1 RESPONSES CORRECTLY REPRODUCED IN RECALL AND RELEARNING

Measure	Cond.	Paradigm			
		A-B, A-B'[a]		A-B, A-C[a]	
		Mean	*SD*	Mean	*SD*
MMFR Groups					
MMFR	Work	3.06	1.20	2.12	.86
	Rest	3.81	.39	3.75	.39
Conventional recall	Work	2.31	1.53	1.69	.76
	Rest	3.43	.66	3.69	.66
10 trials of RL	Work	33.12	5.85	31.75	6.32
	Rest	37.88	2.15	37.31	2.72
Conventional Groups					
Conventional recall	Work	2.50	1.05	1.56	1.00
	Rest	2.94	.50	3.12	.93
10 trials of RL	Work	35.19	3.75	32.94	3.69
	Rest	36.31	2.81	36.75	4.07

[a] Pairs learned by *Ss* in the Rest groups were classified according to treatment for corresponding *Ss* in the Work groups.

than for the A-B, A-C paradigm. Twelve *Ss* contributed 35 intrusions on A-B' pairs whereas 7 *Ss* made 11 such errors on A-C pairs. This difference agrees with that found in other investigations (Barnes & Underwood, 1959; Underwood, 1951; Young, 1955).

Transfer from List 1 to List 2.—To estimate the net amount of transfer from List 1 to List 2, the total numbers of correct responses on the first five trials of acquisition of the two lists were used. Since the fastest learner on List 1 required five trials to reach criterion, all *Ss* could be included in this analysis. For the A-B, A-B' paradigm the mean number of correct responses rose from 7.50 to 12.22; for A-B, A-C there was only a small increase from 6.75 to 8.12. While the net transfer effect is positive in both cases, the gain is substantially larger for A-B' than for A-C. The interaction, Pairs × Lists, is significant beyond the .001 level ($F = 14.73$, $df = 1/31$). In the absence of appropriate controls the effects of learning-to-learn and warm up cannot be separated from the specific transfer effects falling on the two classes of pairs. The large difference in the net amount of gain indicates, however, that the specific transfer effects were positive for A-B' and negative for A-C.

MMFR test.—The mean numbers of correct responses on the MMFR test are shown in Table 1. The scores are based on responses given to the appropriate stimulus terms. For the Rest group there is only a negligible decline in the recall of List 1. The Work group shows losses for both kinds of List 1 pairs but the amount forgotten is clearly greater after interpolation of A-C than of A-B'. Recall

of List 2 was nearly perfect—3.94 for A-B′ and 3.88 for A-C.

In the statistical analysis of the results, the scores of corresponding *Ss* in the Work and Rest groups were treated as paired replicates. Since the direction of the differences was predicted, one-tailed tests of significance were used. The overall amount recalled is significantly greater under Rest than Work ($t = 4.67$, $df = 15$, $P < .001$). For purposes of evaluating the differential effects of the conditions of transfer on the amount of RI, the difference between the recall scores for the two sets of pairs was determined for each *S*. The mean difference score is significantly higher for the Work than the Rest group ($t = 2.56$, $.01 < P < .02$). Thus, the reduction in the availability of List 1 responses is reliably greater for the A-B, A-C than the A-B, A-B′ paradigm.

There were only a few scattered instances of List 1 responses given to an incorrect stimulus on the MMFR test. Inclusion of these responses raises the total number of responses recalled by 2 and 3 for the A-B, A-B′ and A-B, A-C conditions, respectively. For the rest group, there is an increase of 2 on each type of pair. Thus the picture remains virtually unchanged whether or not response to the appropriate stimulus is used as a criterion of correct recall.

Order of recall in MMFR.—The numbers of cases in which *Ss* recalled both responses to a common stimulus term were 48 for A-B, A-B′ and 33 for A-B, A-C. The percentages of cases in which List 1 responses were recalled first were 45.8 and 45.4, respectively, i.e., the frequencies are close to chance in both cases. The corresponding percentages obtained by Barnes and Underwood after 20 trials of interpolated learning were 53 and 43. Since all percentages are near chance, the differences between the two experiments may be considered minor.

Conventional recall of List 1.—Table 1 shows the mean numbers of items recalled by the Conventional groups on the paced test of anticipation. The total recall scores of the Work group are significantly lower than those of the Rest group ($t = 3.51$, $P < .01$). The amount of interference is substantially greater for the A-B, A-C than the A-B, A-B′ paradigm, and the difference between the two sets of pairs is again significantly larger for the Work than the Rest group ($t = 2.38$, $.01 < P < .02$). There was only one interlist intrusion to an appropriate stimulus term, which was a substitution of B′ for B.

Comparison of MMFR and conventional recall.—The overall level of recall is significantly higher in MMFR than in conventional anticipation ($F = 10.36$, $df = 1/15$, $P < .01$, after a Freeman-Tukey square-root transformation). However, the amounts of RI are quite similar under the two conditions of testing. The percentages of RI in MMFR are 19.7 for A-B, A-B′ and 43.5 for A-B, A-C. On the conventional test the corresponding percentages are 14.9 and 50.0. The Rest-Work differences interact with the method of testing neither for total recall scores nor for the differences between types of pairs.

Effect of MMFR on conventional recall.—Since all groups relearned List 1 by the anticipation method, it is possible to evaluate the effects of an MMFR test on subsequent conventional recall. It must be recognized that the time between the last trial of original learning and the first trial of relearning was slightly longer for the MMFR groups than the Conventional groups because of the interpolated MMFR test. However, this differ-

ence is minor relative to the total length of the retention interval. As Table 1 shows, the performance of the MMFR Rest group exceeds that of the Conventional Rest group on the first trial of relearning whereas the differences between the two Work groups are small and not consistent. Whatever opportunity for rehearsal is provided by the MMFR test appears to have a beneficial effect only for the Rest group. However, the resulting increase in RI is not significant as evaluated by the interaction of the Work-Rest differences with the conditions of testing.

The number of appropriate interlist intrusions also shows some increase after the MMFR test. Whereas there was only one intrusion of a B′ item for the Conventional Work group, the MMFR Work group gave four B′ responses and two C responses on the first relearning trial. The interpolation of an MMFR test appears to reduce the differentiation between lists.

Relearning.—The mean numbers of correct responses in 10 trials of relearning are shown in Table 1. The Conventional groups will be considered first. The difference between the Work group and the Rest group is significant ($t = 2.38$, $.01 < P < .02$), i.e., there is reliable RI when the total performance in relearning is considered. However, the difference between the amounts of RI under the two conditions of interpolation falls short of significance ($t = 1.64$). As Fig. 1 shows, the level of RI for A-B, A-C declines steeply and converges on that for A-B, A-B′.

The overall difference between the Work group and the Rest group is also significant for relearning after MMFR ($t = 3.30$, $P < .01$). The amount of RI does not differ significantly for the two paradigms ($t = .54$),

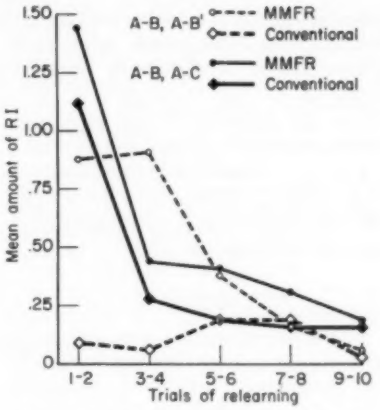

FIG. 1. Amount of RI in relearning as a function of transfer paradigm and condition of testing. (All *Ss* relearned by the method of anticipation. An MMFR test preceded relearning for the MMFR groups but not for the Conventional groups.)

and the loss scores again converge during relearning (Fig. 1).

Comparison of the temporal trends in RI shows that interpolation of an MMFR test produces a pronounced increase in RI for A-B, A-B′ in the early stages of relearning, and a smaller increase for A-B, A-C. As a result, the initial advantage of A-B, A-B′ is considerably greater under the conventional treatment than after MMFR: in fact, the relationship between the paradigms is temporarily reversed in the latter case. A trend analysis of the differences between the two sets of pairs yields a significant interaction of Conditions (Work vs. Rest) with the Method of Testing ($F = 4.77$, $df = 1/15$, $.02 < P < .05$). Thus, the difference between the slopes of the two RI functions is reduced by the interpolation of an MMFR test.

The effects of MMFR on relearning are reflected in the frequencies of interlist intrusions given to the appropriate stimulus terms. These fre-

quencies and the numbers of *Ss* (N) contributing them were as follows: Conventional Work group—8 intrusions of B′ ($N = 5$) and 4 of C ($N = 2$); MMFR Work group—21 intrusions of B′ ($N = 12$) and 5 of C ($N = 4$). There are more intrusions of B′ than C responses in both cases, but this difference is substantially increased following MMFR.

Discussion

The results obtained with mixed lists fully confirm those reported by Barnes and Underwood (1959) for homogeneous lists. The reduction in availability of List 1 associations as measured by MMFR is clearly and significantly greater for the A-B, A-C than the A-B, A-B′ paradigm. The relative amounts of interference observed in the two experiments are in substantial agreement. A direct comparison can be made of the percentages that List 1 responses were of all the responses given in MMFR under each condition of interpolation. For A-B, A-B′ this percentage was 47.0 in the study of Barnes and Underwood, and 43.8 in the present experiment; for A-B, A-C the corresponding percentages are 35.8 and 35.4.

The consistency of the results obtained with homogeneous and with mixed lists indicates that the differences obtained in MMFR reflect primarily the strength of first-list associations rather than the influence of response sets at the time of recall. This finding parallels that of Twedt and Underwood (1959) that homogeneous lists and mixed lists yield equivalent measures of differential transfer effects.

With response set eliminated as a major determinant of performance in MMFR, the present results give further strong support to the hypothesis that first-list associations are extinguished or unlearned when intertask transfer is negative. The data are also consistent with the hypothesis of response mediation under conditions of positive transfer. When response similarity is high, the acquisition of List 2 associations is rapid, and List 1 associations are maintained at relatively high strength. However, there is clearly some decline in the strength of A-B after interpolation of A-B′. As Barnes and Underwood have pointed out, the possibility that there is some extinction of A-B during acquisition of A-B′ must remain open.

The MMFR procedure is designed to measure associative strength independently of the effects of list differentiation and response competition. Nevertheless the MMFR and Conventional groups yield comparable measures of RI and are equally sensitive to the differences between the paradigms of transfer. It appears that in conventional recall the amounts of RI observed immediately after the end of interpolated learning are largely a function of associative strength rather than degree of list differentiation and response competition. The relationship between the measures of RI obtained in MMFR and conventional recall is likely to change, however, with the progressive decline of list differentiation as a function of time.

The rise in intrusions following MMFR indicates a systematic decrease in list differentiation after a test in which *Ss* may give the two responses to a common stimulus term in either order. The fact that the increase in intrusions is greater for A-B, A-B′ than for A-B, A-C lends some indirect support to the view that mediation occurs in the former case. The associative connection between highly similar responses may be assumed to be bidirectional, i.e., the pre-experimental probability of B′-B should be equal to that of B-B′. Both sequences occur with approximately equal frequency in MMFR. Reversal of the order of responses in MMFR appears to be conducive to competition between the two alternative sequences and results in persistent interlist intrusions and RI in relearning. There are presumably no such competing sequences in the case of the A-B, A-C paradigm. These effects of MMFR on relearning add to the evidence for the lack of independence between successive tests of retention.

Summary

This study is concerned with the changes in the strength of first-list associations during the acquisition of a second list. A mixed-list design was used so that for half the syllable-adjective pairs in the two lists the paradigm of transfer was A-B, A-B′, and for the other half, A-B, A-C. List 1 was learned to one perfect recitation and List 2 was practiced for 20 trials. To determine the availability of List 1 and List 2 associations at the end of interpolated learning, a procedure devised by Barnes and Underwood (1959) was followed and *Ss* were required to give both responses to each stimulus in an unpaced test of recall (MMFR). A control group learned and recalled a single list. In a parallel experiment retention for List 1 was tested by the conventional method of anticipation.

In agreement with the results obtained by Barnes and Underwood for homogeneous lists, the MMFR test shows a significantly greater reduction in the availability of List 1 responses for the A-B, A-C than the A-B, A-B′ paradigm. Recall of List 2 was nearly perfect. These findings support the hypothesis of unlearning and make it unlikely that the differences in the availability of List 1 associations are a function of response sets characteristic of the two paradigms. The conventional anticipation method yields measures of RI which closely parallel those obtained in MMFR. It is concluded that under both conditions of measurement the amount of RI immediately after interpolated learning is determined largely if not entirely by the availability of List 1 associations.

References

Barnes, J. M., & Underwood, B. J. "Fate" of first-list associations in transfer theory. *J. exp. Psychol.*, 1959, **58,** 97–105.

Briggs, G. E. Acquisition, extinction, and recovery functions in retroactive inhibition. *J. exp. Psychol.*, 1954, **47,** 285–293.

Briggs, G. E., Thompson, R. F., & Brogden, W. J. Retention functions in reproductive inhibition. *J. exp. Psychol.*, 1954, **48,** 419–428.

Glaze, J. A. The association value of nonsense syllables. *J. genet Psychol.*, 1928, **35,** 255–269.

Haagen, C. H. Synonymity, vividness, familiarity, and association-value ratings for 400 pairs of common adjectives. *J. Psychol.*, 1949, **30,** 185–200.

Melton, A. W. Comments on Professor Postman's paper. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior.* New York: McGraw-Hill, 1961.

Melton, A. W., & Irwin, J. McQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53,** 175–203.

Melton, A. W., & Von Lackum, W. J. Retroactive and proactive inhibition in retention: Evidence for a two-factor theory of retroactive inhibition *Amer. J. Psychol.*, 1941, **54,** 157–173.

Newton, J. M., & Wickens, D. D. Retroactive inhibition as a function of the temporal position of interpolated learning. *J. exp. Psychol.*, 1956, **51,** 149–154.

Osgood, C. E. Meaningful similarity and interference in learning. *J. exp. Psychol.*, 1946, **36,** 277–301.

Osgood, C. E. An investigation into the causes of retroactive inhibition. *J. exp. Psychol.*, 1948, **38,** 132–154.

Postman, L. The present status of interference theory. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior.* New York: McGraw-Hill, 1961.

Postman, L., & Riley, D. A. Degree of learning and interserial interference in retention. *U. Calif. Publ. Psychol.*, 1959, **8,** 271–396.

Twedt, H. M., & Underwood, B. J. Mixed vs. unmixed lists in transfer studies. *J. exp. Psychol.*, 1959, **58,** 111–116.

Underwood, B. J. Proactive and retroactive inhibition after five and forty-eight hours. *J. exp. Psychol.*, 1948, **38,** 29–38. (a)

Underwood, B. J. "Spontaneous" recovery of verbal associations. *J. exp. Psychol.*, 1948, **38,** 429–439. (b)

Underwood, B. J. Associative transfer in verbal learning as a function of response similarity and degree of first-list learning. *J. exp. Psychol.*, 1951, **42,** 44–53.

Young, R. K. Retroactive and proactive effects under varying conditions of response similarity. *J. exp. Psychol.*, 1955, **50,** 113–119.

(Received September 7. 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 388–392

EFFECTS OF NONREINFORCEMENT ON SUBSEQUENT REINFORCED RUNNING BEHAVIOR [1]

SHELBY J. HARRIS,[2] M. GLENN SMITH,[3] AND SOLOMON WEINSTOCK [4]

Lehigh University

Bush and Mosteller (1955, p. 89) present a descriptive model for response acquisition under partial reinforcement which is formally identical with models arising from statistical learning theory. They assume two linear operators on probability which are applied independently, one on reinforced and the other on nonreinforced trials. The model yields the result that asymptotic probability of responding is directly related to the proportion of reinforced trials. This result may also be shown to hold for the asymptotic speed of running in an alleyway. The latter was directly contradicted by the finding by Notterman (1951), Weinstock (1958), and others, that continuously reinforced *S*s reached lower terminal speeds of running than partially reinforced *S*s.

Weinstock (1953, 1958) hypothesized that the effect of nonreinforced trials was to provide an opportunity for nonfunctional response components in the response chain to be eliminated. Elimination of nonfunctional movements should result in an increase in the value of the limit-point of the reinforced trial operator. Weinstock (1953) proposed a modification of the original model in which the limit-point of the reinforced trial operator is assumed to be an increasing function of the number of nonreinforced trials.

In the present experiment two groups of rats were trained to run in an alleyway under 50% and 100% reinforcement, respectively. Half of each group then received extinction. Finally, all *S*s were given 100% reinforcement. In the last phase of the experiment only the reinforced trial operator is applicable and the changes induced in the limit-point should be clearly exhibited.

It was expected that both the partial reinforcement and extinction variables would lead to faster running and that an interaction between the two variables would be found. It was also expected that the 50% group which did not get extinction would run faster when switched to 100% reinforcement since the nonreinforced trial operator would no longer be applied. Finally, it was hoped that the differences induced in speed of running would correspond to differences in within-trial variability. Specifically, it was thought that faster running was due to conditioning of a response chain that contained fewer nonfunctional movements. As *S* traversed the alleyway, the successive placements of its left rear paw were recorded. More general information about the effect of reinforcement schedules on the variability of running behavior was also sought.

METHOD

Subjects.—The *S*s were 40 experimentally naive female albino rats, 105 to 135 days old at the beginning of the experiment.

Apparatus.—The apparatus was a straight alley, 10 in. wide and 36 in. long, with a 10 × 18 in. goal box and a start box 5 in. wide

[1] This study was supported in part by Research Grant M-2286 from the National Institute of Mental Health, United States Public Health Service.

[2] Now at Wells College.

[3] Now at Tufts University.

[4] Now at Brooklyn College.

and 10 in. long. The height of all sections was 8 in. An entranceway, 5 in. wide and 4 in. long, separated the alley from the goal box. Guillotine doors were used in the start box and in the goal box. A 5 × 7 in. vestibule in the far right corner of the goal box contained a drinking tube. On nonreinforced trials a sliding door, which was not visible before *S* interrupted the last photocell beam, prevented access to the drinking tube. Latency and total running time were measured using Standard Electric .01-sec. timers operated by photocells located 2 in. from the start box and 9 in. inside the goal box, respectively.

Records of the path taken by *S* were obtained by means of a 20-channel "contact" recorder in which *S*'s left rear paw served to key a thyratron circuit when the paw made contact with the floor. This was accomplished by constructing the floor of tinned copper bus bars embedded in insulated wood at ¼-in. intervals. When the paw, which was coated with silver paint, made contact with adjacent bars a thyratron circuit was completed. Resistance in the other paws was too high to complete the circuit.

The runway floor was divided into a checkerboard of ¾ × 2 in. sections. Contact within one of these sections activated two electrodes of a 20-electrode paper-burn recorder. One of the electrodes recorded the position of the paw across the runway and the other the position down the length of the runway. Eleven ¾-in. and one 1½-in. segments, adjacent to the right hand wall, made up the width of the alleyway. Six 2-in. sections made up the first 12 in. of the length of the runway. Two repetitions of this matrix were used for the remaining portion of the runway proper and a third repetition permitted recording in the goal box. Two photocells, one located 18 in. from the start box and the other 9 in. inside the goal box, operated two of the remaining electrodes of the recorder to aid in determining when *S* moved from one section to the next.

Procedure.—Prior to the experiment, *S*s were handled and habituated to 22 hr. water deprivation for 10 days. Throughout the experiment two trials per day with an intertrial interval of about 45 min. were run under 22 hr. water deprivation. Reinforcement consisted of 10 sec. of drinking. On nonreinforced trials *S* was confined to the goal box for 30 sec.

Before each trial *S*'s left rear paw was coated with silver paint. After a 30-sec. interval to allow the paint to dry, *S* was placed in the start box facing the door. Closing the cover depressed a microswitch which activated a 3-sec. thermo-delay relay. When the relay closed a solenoid was energized which served to release a lever which was held by stretched rubber bands and both timers were started. Movement of the lever raised the start-box door.

Two *E*s (SJH and SW) ran half of the *S*s each. The third *E* coated *S*'s paw with paint and recorded the data.

During the first four trials of the experiment, *S* was removed from the apparatus if it failed to leave the start box within 20 min. or spent more than 20 min. in the runway or in the goal box without being reinforced. After the fourth trial the removal criteria were reduced to 5 min. Failure to run on three successive trials was used as a criterion for eliminating *S* from the experiment.

All *S*s were reinforced on each of the first 12 trials. On the basis of running speed data from the early trials *S*s were assigned to four groups in such a way that the group means were equal.[5] Group CE remained on 100% reinforcement during the acquisition period and was then given 24 extinction trials. Group CC remained on 100% reinforcement, both during acquisition and during the subsequent 24 trials. Group PE received 50% reinforcement during acquisition and then received 24 extinction trials, while Group PP received 50% reinforcement, both during acquisition and the following 24 trials. For *S*s receiving 50% reinforcement 6 trials on which reinforcement was to be given were selected randomly from each block of 12 trials.

Mean reciprocal running times were computed for each block of six trials during the course of acquisition. The decision to terminate acquisition was based on inspection of these data. When it appeared *S*s' running times were stable, an additional 12 trials were run and the acquisition period terminated. As a result of this procedure a total of 96 acquisition trials were run. Prior to the extinction period all *S*s received 2 reinforced trials. A reacquisition period of 30 trials followed the 24 extinction trials. All *S*s received 100% reinforcement during reacquisition.

Results

Latency and running time were converted to reciprocals throughout.

[5] One *S* was discarded after Trial 9 in accordance with the criteria previously described. Hence, for Group PE $N = 9$, while for the other groups $N = 10$.

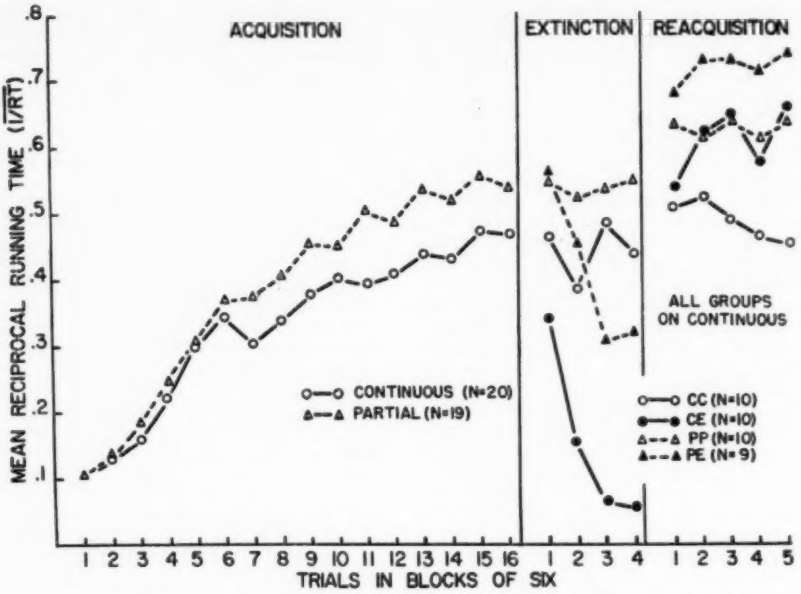

FIG. 1. Mean reciprocal running time plotted as a function of trials in blocks of six.

A record of successive placements of *S*'s left rear paw was obtained on each trial. The paw placements were plotted on a scale drawing of the runway and connected by straight lines yielding an approximate path for each trial. The points of intersection of the path with lines perpendicular to the side walls at distances of 7, 14, 21, and 28 in., respectively, were recorded. Thus, the lateral position of *S* (i.e., across the width of the runway) at each of the four points down the runway could be estimated. From these data a number of measures involving spatial location could be computed for each *S* for any block of trials: (*a*) the variability of lateral position across a block of trials at each of the four points, (*b*) the mean lateral position across all four points and across trials, (*c*) the variability of this mean position across trials, (*d*) the mean across trials of the variance within a trial computed about the mean position, and (*e*) the sum within each trial of the absolute deviations from a straight line connecting the start box and goal box and the average of these sums over trials.

In Fig. 1 the reciprocal running time measure is plotted against blocks of six trials for all three phases of the experiment.[6] Reacquisition occurred much more rapidly than original acquisition. Group CE went from a level of responding below that of the first block of acquisition trials to a level about that of Group CC within the first block of reacquisition trials. The major portion of the change in responding for Group CE took place in the first three reacquisition trials.

An analysis of variance for the last 18 trials of acquisition yielded a non-

[6] The two trials at the end of acquisition on which all *S*s received reinforcement are omitted from Fig. 1.

significant F ($F = 2.60$, $df = 1/35$, $P > .05$) for the difference between the continuous and partial reinforcement groups. The difference between Groups CC and PP continued to be nonsignificant for the 24 trials of the extinction phase, while the difference between Groups CE and PE yielded an F of 20.32 with 1/15 df, which is significant at the 5% level.

The results of an analysis of variance based on the last 12 trials of the reacquisition phase are summarized in Table 1. Since 7 Fs were computed, Hartley's F maximum test was used to adjust the experiment-wise error rate to 5%. Pearson and Hartley (1954, Table 19) give the 5% value as slightly less than 8.21. Thus, with the experiment-wise error rate controlled at 5% the effects of the reinforcement and extinction variables would be declared significant while the interaction between Es and extinction would be declared nonsignificant. It should be noted that the size of the interaction was due to a larger difference between extinction and no-extinction groups for one E (SJH) than for the other (SW). The direction of the difference was the same for both Es.

Group PP showed an increase in speed of running after being shifted to continuous reinforcement.

TABLE 1

ANALYSIS OF VARIANCE FOR THE REACQUISITION PHASE

Source	df	MS	F
Es	1	.0077	—
Continuous vs. Partial	1	.1811	12.32*
Extinction vs. No-Ext.	1	.1819	12.37*
Es × C-P	1	.0154	1.05
Es × E-NE	1	.0729	4.96*
E-NE × C-P	1	.0097	—
E-NE × C-P × Es	1	.0222	1.51
Within (error)	31	.0147	

* $P < .05$.

The group's mean level during the extinction and the reacquisition phases was 0.54 and 0.63, respectively. A t for paired measures, based on the last 18 trials of each of the two phases, yielded a value of 3.04 with 9 df, which is significant at the 5% level.

Sixteen analyses employing the measures based on S's path were performed, 8 at the end of acquisition and 8 at the end of reacquisition. With one exception they proved to be nonsignificant. At the end of acquisition the partial reinforcement group showed "significantly" more variability in lateral position at a distance of 14 in. from the start box. While no method of adjusting these results to a 5% experiment-wise error rate exists, common sense suggests that the result not be declared significant.

Mean reciprocal running times for the last 12 trials of initial acquisition were correlated with four measures of Ss' paths on the same trials. The correlations with running speed were: (*a*) mean position of S in the lateral dimension (i.e. width) of the runway, −.09; (*b*) average within-trial variance of lateral position computed about the mean position, −.24; (*c*) variance of the mean lateral position across trials, .03; and (*d*) average sum of absolute deviations from a straight line connecting the start and goal boxes, −.20.

DISCUSSION

The major findings involving the time data of this experiment were that nonreinforced trials, administered either in an extinction or partial reinforcement procedure, led to a change in the terminal level of responding under subsequent continuous reinforcement. There was no sign of a decrease in the higher level of responding induced by nonreinforcement, which suggests that a nonreversible change in the reinforced trial operator has occurred. Although appropriate

controls were lacking, the results also suggest that the larger the number of nonreinforced trials, the greater the effect induced on the reinforced trial operator. To this extent the qualitative predictions stemming from a modification of the Bush and Mosteller model received support. It is difficult to assess the failure to find the predicted interaction between the extinction and proportion of reinforcement variables. Ideally a power analysis should be performed. However, this is impossible without a good estimate of the expected magnitude of the interaction, which depends critically on the unknown parameters of the various learning operators.

The problem of providing a theoretical underpining for the induced changes in the reinforced trial operator remains. One attempt in this direction (Weinstock, 1953, 1958) leads to the prediction that the changes in speed of running should be reflected in a reduction of within-trial variability. This prediction received no support from the present results.

In general, the spatial location measures were not related either to the independent variables investigated or to speed of running. The question of the reliability of the measures arises. To obtain a rough answer, values for each of the four major measures for Trials 86 to 91 were correlated with those for Trials 92 to 97. Values of .77, .66, .86, and −.19 were obtained for the absolute deviations from the midline, the variance within a trial, the mean position, and the variance of the mean position, respectively. For purposes of comparison the correlation coefficient was also computed for the running time measure for the same blocks of trials and a value of .83 was obtained. With the exception of the variance of the mean position, it seems that the measures taken were sufficiently reliable.

Summary

Forty albino rats were given two trials per day, separated by 45 min., in a straight alley under 22 hr. of water deprivation. Half of the *S*s received 100% and half 50% reinforcement for 96 trials of acquisition. The two groups were further halved with half of each receiving 24 extinction trials and the other half continuing on their reinforcement schedules. All *S*s then received 30 trials under 100% reinforcement. In addition to latency and running time measures, detailed recording was made, by use of a special "contact" recorder, of the path taken by *S*.

The following results were obtained. At the end of acquisition 50% *S*s were running more rapidly than 100% *S*s but the difference was not statistically significant. During the last 30 trials of the experiment *S*s who had had 50% reinforcement were running more rapidly than *S*s who had had 100% reinforcement, and *S*s who had had extinction were running more rapidly than those who had not. The 50% reinforcement *S*s ran more rapidly after being shifted to 100%. These differences were statistically significant. The interaction between the extinction and proportion of reinforcement variables was not significant.

In general the various measures of spatial location and variability which were taken showed no relation to the independent variables and quite low correlations with speed of running.

References

Bush, R. R., & Mosteller, F. *Stochastic models for learning.* N. Y.: Wiley, 1955.

Notterman, J. M. A study of some relations among aperiodic reinforcement, discrimination training, and secondary reinforcement. *J. exp. Psychol.*, 1951, **42**, 273–291.

Pearson, E. S., & Hartley, H. O. *Biometrika tables for statisticians.* Cambridge: Cambridge Univer. Press, 1954.

Weinstock, S. Acquisition and extinction of a partially reinforced running response at a 24-hour intertrial interval. Unpublished doctoral dissertation, Indiana University, 1953.

Weinstock, S. Acquisition and extinction of a partially reinforced running response at a 24-hour intertrial interval. *J. exp. Psychol.*, 1958, **56**, 151–158.

(Received September 8, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 393–399

EFFECTS OF ORIENTING TASK, PRACTICE, AND INCENTIVE ON SIMULTANEOUS INCIDENTAL AND INTENTIONAL LEARNING[1]

ARNOLD MECHANIC[2]

Northwestern University

In the Type I incidental learning situation, *S* is given no instructions to learn but is later tested for the materials to which he was exposed. In the Type II situation, *S* is exposed to two sets of materials, instructed to learn only one of the sets, and is later tested for the materials which he was not instructed to learn. The finding that there is less incidental learning in the Type II situation than in Type I has been attributed to competition from the intentional task for the limited amount of exposure time which is available to *S* (Mechanic, 1962). It could therefore be expected that certain intentional learning variables would adversely affect Type II incidental learning. Any experimental variation which focuses *S* on the intentional items should serve to reduce the proportion of the total exposure time available for responding to the incidental items. Examples of such variables would be strength of the incentive motivating *S* to learn, and difficulty of the intentional items.

Bahrick (1954) and Bahrick, Fitts, and Rankin (1952), using nonverbal materials, found that the amount of learning irrelevant to a set is inversely related to the strength of the incentive determining that set. Mechanic (1962), however, found that incidental learning of verbal items did not vary as a function of the difficulty of the concurrently learned intentional items. The divergence between these findings, with regard to the task-competition hypothesis, becomes explicable when we consider the different orienting tasks which *S*s were required to perform. Mechanic's *S*s were required to pronounce the incidental items in order to rate phonetic similarity. Thus, competition from the intentional items (for the available exposure time) could not prevent *S*s from responding to the incidental items. In Bahrick's study, on the other hand, there was no specific orienting task which required *S* to respond to the incidental stimuli. The *S* could, for example, learn geometric forms intentionally without necessarily responding to their incidental colors, and large amounts of competition from the intentional task could result in a failure to respond to the incidental stimuli.

If the above explanation is appropriate, it would be expected that the incentive for intentional learning would have little or no effect on concurrent incidental learning when the orienting task required that *S* respond to the incidental stimuli. This experiment is an attempt to clarify the joint effects of orienting task and intentional learning variables on Type II incidental learning. By varying the degree of responding to the incidental items required by the orienting task,

[1] This research was conducted during the writer's tenure as a National Science Foundation Postdoctoral Fellow. The writer is indebted to B. J. Underwood for his valuable suggestions and critical reading of the manuscript.

[2] Now at Alameda State College, Hayward, California.

it should be possible to get a clearer picture of how intentional learning variables influence Type II incidental learning. Because the intentional learning variable that has been chosen is incentive, it will, at the same time, be possible to assess the effects of incentive upon the intentional rote learning of verbal materials.

METHOD

General design.—Three orienting tasks were selected so as to differ in the degree of responding that they required to the incidental stimuli. Two levels of incentive were used. In addition, two degrees of practice (number of presentations) were used in order to obtain greater generality in the evaluation of the incentive and orienting task variables, and in order to study possible interactions between these variables and practice. Twelve groups of 20 *Ss* each were used in a standard $3 \times 2 \times 2$ factorial design.

All *Ss* were given a list made up of 12 pairs of trigrams. They were instructed to learn one member from each pair and were discouraged from attempting to learn the other member of the pair. Different groups of *Ss* were exposed to the items by means of the three orienting tasks described below. The *Ss* were told that the two members of each pair were syllables from two different primitive languages. The first members of each pair were said to come from one language while the second members of the pairs came from the other language. In addition to the orienting task instructions, *Ss* were told that we also wanted to study the ease of learning of one primitive language under conditions of distraction from a second language. In this way, instructions to learn were given for only one member of each pair (intentional learning of one member and incidental learning of the other). After the list presentations, a test of free recall was given for all the items. Each *S* received two learning scores, one being the number of correctly reproduced intentional items and the other being the number of correctly reproduced incidental items.

Orienting tasks.—Three orienting tasks were devised so as to produce differing numbers of "pronouncing" responses to the incidental items. By a pronouncing response, reference is made to a hypothetical response by *S* so that he reacts to the trigram as a single pronunciable unit. The operations involved in the three orienting tasks should serve to further clarify this definition. The choice of the pronouncing response as an important response for rote learning is consistent with the recent conclusion of Underwood and Schulz (1960) that, "*emitted* frequency is of fundamental importance in the integration of response elements [italics mine]" (p. 291). The orienting task devised to produce maximal responding to the incidental items will be called Phonetic Similarity ratings (PS). To PS *Ss*, the experiment was introduced as a test of the notion that words meaning the same thing in different languages tend to sound more alike than words with different meanings. The *Ss* were asked to judge the phonetic similarity between the members of each pair of trigrams. The *Ss* were required to pronounce the members of each pair to themselves and to rate their phonetic similarity on a five-point scale. Following these instructions, learning instructions for only one set of items were given as described above.

It may be noted that this orienting task required that *S* respond to the incidental items as single pronunciable units. Another orienting task was designed in order to minimize the probability of *S* responding to the incidental items in this manner. This minimal responding situation will be referred to as Letter Cancellation (LC). The LC *Ss* were told that we were interested in the effects of quick changes of set upon letter cancellation performance. They were required to cancel out the vowels on the top language (pair member) and cancel out the consonants on the bottom language. All letters other than A, E, I, O, and U were considered consonants. The *Ss* were asked to work carefully and were not permitted to erase after having marked a letter. It is clear that this orienting task exposes *S* to the incidental stimuli without necessarily requiring that he respond to the items as pronunciable units.

Having defined extreme points on a hypothetical dimension, an intermediate orienting task was selected, which will be called Extrasensory Perception (ESP). The ESP *Ss* were told that of the 144 possible pairings of the items from the two languages, 50 had been selected randomly and printed on cards. These 50 cards were said to be in an envelope that was left on the desk in front of the room. The *Ss* were told that we wanted to check the contention of previous investigators that people could guess better than would be expected by chance in such a situation. They were required to look at each pair of syllables and then rate the likelihood of its being represented in the envelope. As with the

phonetic similarity judgments, the ratings of likelihood were on a five-point scale. This kind of orienting task was chosen because it seemed to offer some probability of *S* responding to the incidental items as pronunciable units. At the same time, such responses were not required as they were in the PS situation, and should be of lower frequency for the ESP group. However, the ESP group could clearly be expected to make more pronouncing responses than the LC group.

It should be evident that the amount of time required to perform each of the three orienting tasks will vary, and that effective exposure times for learning will vary concomitantly. Although there are no quantitative data on this issue, it is apparent from observations made during the experiment that PS ratings take the most time while LC takes the least time. Because both intentional and incidental scores were obtained from each *S*, it will be possible to gauge the incidental scores for each orienting task by the corresponding intentional scores for that condition. The intentional scores may be thought of as an index of the time available for learning to take place under each respective orienting task.

Incentive conditions.—The *S*s were introductory psychology students at Northwestern University who met a course requirement by serving as *S*s. In addition, "incentive" *S*s were offered the possibility of one "bonus hour" toward the required total of 10 hr. They were told that an extra hour of credit would be given to all *S*s who learned more of the intentional syllables than the average number learned by the group. It was added that they would be eligible for the bonus only if their performance on the other task (orienting) met certain standards of acceptability. The LC and PS *S*s were told that they would be eligible for the bonus only if their performances met minimal standards of accuracy. The ESP *S*s were told that they would not be eligible if their guesses were in the bottom 10% of the group. It was made clear, however, that half of the *S*s in the group would win bonus hours and would be informed of these awards promptly.

The "nonincentive" *S*s were not offered bonus hours, but were otherwise read instructions identical to those of their incentive counterparts. Incentive *S*s were questioned about the effects of the incentive during a postexperimental inquiry. Those who stated that they did not need the bonus hour, and that they did not try any harder as a result of the incentive, were discarded and replaced. The positive statements of many *S*s would seem to indicate that a bonus hour has incentive value in the same sense as do the small monetary bonuses paid by Bahrick (1954).

Practice conditions.—The list of 12 pairs was exposed for either two or five presentations. Successive presentations of the list involved different random orders for the set of 12 intentional items and the set of 12 incidental items. As a result, the pairings of individual items from the two sets were randomly varied from presentation to presentation. This was done to prevent the learning of incidental items through the mediation of the intentional members of the pairs.

Design details.—The 12-pair lists were constructed from two separate 12-unit trigram lists. The two 12-unit lists were equated for meaningfulness and were made up of pronunciable high-frequency items. The two lists and their method of construction have been presented elsewhere (Mechanic, 1962). For clarity of exposition, henceforth these 12-unit lists will be referred to as *sets* and the term *list* will refer to the 12-pair lists actually presented to the *S*s. Each of the 12 pairs in a list was arranged vertically with one member in the top position and the other member directly below it. This arrangement was used to prevent the left to right chaining which might result from the reading habits of *S*s. The top members of the pairs came from one of the stimulus sets while the bottom members came from the other set. Through presentation of the list during training, both sets of items were presented concurrently to *S*. The *S*s were, of course, instructed to learn only one of the two sets.

Five random orders were prepared for each of the sets. This was done to obtain different orders of the items with successive presentations of the list. Additionally this insured that the pairings of individual items from the two sets were randomly varied from trial to trial. For each of the five orders of presentation, two variations were prepared. These reversed the top and bottom member positions but were otherwise identical.

The *S*s were given a number of sheets of paper, each covered with a strip of cardboard in which a cut-out window allowed exposure of one pair of syllables at a time in accordance with *E*'s instructions. The pairs of syllables on each sheet were numbered 1 through 12. When *E* called out the number 1, *S* moved the window to expose the first pair of syllables and responded to the items in accordance with the requirements of the orienting task. The procedure was repeated for each successive

TABLE 1

AVERAGE NUMBERS OF INTENTIONAL AND INCIDENTAL ITEMS RECALLED BY THE VARIOUS GROUPS

Type of Item and Incentive	Two Presentations						Five Presentations					
	Task LC		Task ESP		Task PS		Task LC		Task ESP		Task PS	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Incidental Items												
Normal	.70	.56	1.90	1.38	2.75	1.48	1.85	1.77	2.65	1.93	4.50	2.17
Augmented	.65	1.11	1.30	1.05	2.75	2.16	1.75	1.58	2.35	2.24	4.45	1.66
Total	.68		1.60		2.75		1.80		2.50		4.48	
Intentional Items												
Normal	7.55	2.16	6.75	2.62	5.65	1.53	10.25	1.97	8.35	1.96	8.15	2.24
Augmented	7.40	1.50	7.05	2.16	5.80	2.18	10.30	1.95	8.90	2.30	9.20	1.91
Total	7.48		6.90		5.72		10.28		8.62		8.68	

pair until all 12 pairs on the page had been exposed. The pairs were presented at a 10-sec. rate, with 20 sec. between presentations of separate sheets. A 5-min. test of free recall was given 1 min. after the end of practice. The *Ss* were requested to write down as many of the syllables as they could remember regardless of whether they appeared in the top or bottom positions of the pairs. They were told to write the syllables down in the order they were remembered, and that their scores would depend only upon the number of correct items that they wrote down. When the recall test was completed, a postexperimental inquiry was made to ascertain whether *S* understood the instructions, whether the bonus hour had value as an incentive, and whether an attempt was made to learn the incidental items.

For each of the 12 groups, half the *Ss* received materials with the items from one set in the top positions of the pairs. For the other half of the group, the otherwise identical pairs were inverted. In addition, half the *Ss* given each variation of materials were asked to learn the top items and the other half were asked to learn the bottom items.

The *Ss* were tested in a large room in groups averaging 14 in number. Conditions were assigned to the groups in a random order for each replication of the basic experimental design. Instructions to learn top items were alternated with instructions to learn bottom items for the different groups serving in the same experimental conditions. The *Ss* within each group were assigned the two different sets of stimulus materials (top-bottom variations) by rotation. The only restrictions on these procedures were to correct for equal *N*s.

Because *Ss* were tested in groups, extra *Ss* were obtained in most of the experimental variations. Those used in the analyses were selected at random from the total number of *Ss* in each variation. These selections were made after discarding *Ss* whose postexperimental reports indicated that they did not understand the instructions, that they tried to learn the incidental items, or that the bonus hour had no value as an incentive. Less than 10% of the *Ss* appearing for the experiment were discarded for one of these reasons.

RESULTS AND DISCUSSION

Table 1 contains means and *SD*s for both the incidental and intentional learning scores for the 12 basic conditions. The incidental and intentional scores for any combination of experimental variables were obtained concurrently from the same *Ss*. Analyses of variance of the incidental and intentional scores were done separately.

Incentive.—If incentive is to be used to study the effects of an intentional learning variable on concurrent incidental learning, it should be possible to demonstrate that incentive is an

effective variable with regard to the intentional items. The data in Table 1 indicate that incentive has no significant effect upon intentional rote verbal learning ($F=1.47$, $df=1/228$). Therefore, it is not surprising that incentive for the intentional items has no effect upon concurrent incidental learning ($F = .69$, $df = 1/228$). It may be concluded that incentive, as here manipulated, is not an effective variable for either intentional or incidental learning in the Type II verbal learning situation. For the purpose of evaluating other experimental variables, the normal and augmented incentive conditions may be regarded as replications of the experimental design.

Degree of practice.—Both intentional and Type II incidental learning increase significantly as a function of number of presentations ($F = 86.83$ and $F=32.23$, respectively, $df=1/228$, $P < .001$ in both cases). These data confirm earlier findings for the Type II situation (Mechanic, 1962). There are no interactions between practice and the other experimental variables.

Orienting tasks.—Table 1 clearly shows the predicted differences in incidental learning as a function of orienting task. With two incentive levels and two practice levels, there were four replications of the comparison among orienting tasks. In every case, incidental learning increased directly as a function of the hypothesized number of pronouncing responses. There were no reversals of order and the differences within each replication were of substantial magnitude. The overall analysis indicates the highly significant effect of orienting task on incidental learning ($F = 40.05$, $df = 2/228$, $P < .001$).

It could be argued that the effects of orienting task were not due to the hypothetical pronouncing dimension which dictated the choice of the three tasks. Rather, the tasks may have differed in time required to perform them, with concomitant variation in the amount of exposure time available for learning. If this is the case, intentional learning should show the same relation to orienting task as does incidental learning. Previous research indicates that amount of intentional learning may vary with the nature of the orienting task (Postman & Adams, 1956). Again, Table 1 shows four replications of the comparison among tasks. For three of the replications, the order for intentional learning was directly opposite to that for incidental learning. The LC *Ss* showed the greatest learning while the PS *Ss* showed the least learning. The fourth replication (augmented incentive-five presentations) showed a reversal with the PS *Ss* scoring slightly higher than the ESP *Ss*. In spite of this minor reversal, the reduction in intentional learning from LC to ESP to PS is highly significant ($F=13.55$, $df=2/228$, $P<.001$). This would seem to indicate that the predicted differences in incidental learning are not due to the differences in time and effort required by the various orienting tasks. If anything, the predicted differences in incidental learning may be minimized because the orienting tasks requiring more pronouncing responses happen (when gauged by the intentional scores) to be more difficult, or time consuming to perform, than the tasks requiring less pronouncing responses.

It cannot be concluded that these data favor a pronouncing response hypothesis without considering still another alternative interpretation. Perhaps the effects of orienting task are a function of how much time the orienting task requires *S* to devote to the incidental items. If an orienting task requires *S* to spend more time on the incidental items, it is reasonable to assume that less exposure time will be available for the intentional items. This "time" interpretation can explain the finding that from LC to ESP to PS, incidental learning increases while intentional learning decreases. The data are accounted

for without utilization of the hypothetical pronouncing responses which originally dictated the choice of the three orienting tasks. However, the data in Table 1 allow a check on this interpretation by means of a further analysis. It may be noted that at five presentations, the intentional means for ESP (8.35 and 8.90) were approximately equal to the corresponding means for PS (8.15 and 9.20). Whereas intentional learning did not decrease from ESP to PS, incidental learning increased greatly (2.65 and 2.35 vs. 4.50 and 4.45). Even where the intentional scores were not lower, and could not therefore be said to reflect greater time available for incidental learning, incidental learning increased as a function of hypothesized number of pronouncing responses.

At two presentations, it appears that the "time" interpretation is consistent with the data. This interpretation is based upon the distribution of the total exposure time between intentional and incidental learning. It implies that the total of intentional and incidental learning combined should never be greater for a slowly performed orienting task than for a more quickly performed task. It was noted earlier that Task LC was more quickly performed than Task ESP which was more quickly performed than Task PS. Certainly, the time interpretation could not predict that the PS *Ss* would learn more total items (intentional and incidental combined) than the ESP *Ss*. The ESP *Ss* have more time for learning after performing the orienting task. Even if the ESP *Ss*, by spending less time on the orienting task, are learning less incidentally, they should at least make up for this by learning more intentional items in the greater time period remaining. The time interpretation would require that the ESP *Ss* learn no fewer items than the PS *Ss*. The pronouncing response interpretation has no such requirement.

In order to evaluate the above argument statistically, the different incentive groups in Table 1 were combined to give 40 *Ss* for each orienting task at each level of practice. For each *S*, the intentional and incidental scores were combined to give a total learning score. The mean total learning for each orienting task after two presentations is 8.15, 8.50, and 8.48 for LC, ESP, and PS, respectively. Comparable means with five presentations are 12.08, 11.12, and 13.15, respectively. With two presentations, the total learning scores for the three orienting tasks do not differ ($F = .20$, $df = 2/117$). As noted above, the data for two presentations do not distinguish between the pronouncing and time interpretations. While the pronouncing view predicts differences in incidental learning, equality of total learning could still result from differences in amounts of intentional learning permitted by the different orienting tasks.

The data for five presentations argue against a time interpretation. The total learning scores of the three task groups show a significant degree of variation ($F = 4.38$, $df = 2/117$, $.01 < P < .05$). The gap between Groups ESP and PS is significant beyond the .01 level. Thus the differences in incidental learning among the three orienting tasks are not compensated for by reciprocal differences in intentional learning. Clearly, the differences obtained as a function of orienting task are not merely due to different distributions of *S*'s time between the incidental and intentional items. It would seem more reasonable to conclude that the differences in incidental learning among the tasks are a function of the activities required by the different tasks.

Item analysis.—Underlying our view of the incidental verbal-learning situation is the assumption that learning—both intentional and incidental—takes place as a result of pronouncing responses made by *S* to the stimulus items. If these same kinds of responses underlie both intentional and incidental learning, the stimuli facilitating these responses should be the same for both forms of learning. Those stimulus items most frequently learned intentionally should also show most frequent incidental learning. There are data for 24 trigrams for both intentional and incidental learning. The frequencies

with which each item was given as a correct intentional response by 120 *Ss* were correlated with the frequencies with which each item was given as a correct incidental response by the other 120 *Ss*. The product-moment correlation across the 24 items was .81. For purposes of evaluating this correlation, *r*'s were computed for the 24 items between: (*a*) intentional learning at two presentations and intentional learning at five presentations; (*b*) incidental learning at two presentations and incidental learning at five presentations; and (*c*) total learning at two presentations and total learning at five presentations. The *r*'s were .76, .59, and .82, respectively. These three *r*'s were computed as measures of the reliability of item difficulty. Clearly, the item by item correlation between intentional and incidental learning matches the reliability indices of item difficulty. These results are consistent with the belief that the same responses are required in both intentional and incidental learning. In this view, instructions to learn increase learning because they facilitate the performance of the appropriate responses.

Summary

Type II incidental learning refers to the situation in which *S* is exposed to two sets of materials, instructed to learn only one of the sets, and is later tested for the materials which he was not instructed to learn. It may be expected that such learning will vary with: (*a*) the orienting task by which *S* is exposed to the incidental items; (*b*) the incentive for the concurrent intentional items; and (*c*) the number of presentations. Twelve verbal learning groups of 20 *Ss* each were run in a standard 3 × 2 × 2 factorial design, with the variables being orienting task, incentive, and number of presentations, respectively. Separate scores were obtained for intentional and incidental learning from each *S*. Incentive was found to have no effect on either intentional or incidental verbal learning. In agreement with previous results, both kinds of learning increased reliably with number of presentations. Incidental learning increased reliably as a function of the hypothesized number of pronouncing responses required by the orienting task. After considering alternative explanations, it was concluded that the nature of the responses required of *S* by the orienting task are of crucial importance for the amount of incidental learning which occurs.

References

Bahrick, H. P. Incidental learning under two incentive conditions. *J. exp. Psychol.*, 1954, **47**, 170–172.

Bahrick, H. P., Fitts, P. M., & Rankin, R. E. Effect of incentives upon reactions to peripheral stimuli. *J. exp. Psychol.*, 1952, **44**, 400–406.

Mechanic, A. The distribution of recalled items in simultaneous intentional and incidental learning. *J. exp. Psychol.*, 1962, **63**, 593–600.

Postman, L., & Adams, P. A. Studies in incidental learning: IV. The interaction of orienting tasks and stimulus materials. *J. exp. Psychol.*, 1956, **51**, 329–333.

Underwood, B. J., & Schulz, R. W. *Meaningfulness and verbal learning.* Philadelphia: Lippincott, 1960.

(Received September 13, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 400-404

RETROACTIVE INHIBITION OF R–S ASSOCIATIONS[1]

GEOFFREY KEPPEL AND BENTON J. UNDERWOOD

Northwestern University

Consider the S-R association that is formed during the learning of a paired-associate list (A-B). If a second list, bearing an A-C relationship to the first list (i.e., stimuli same, responses different), is interpolated between A-B learning and subsequent A-B recall, the A-B association appears to go through an extinction process (Barnes & Underwood, 1959; Briggs, 1954). This finding can lead to the assumption that extinction will occur whenever an A-B, A-C interference paradigm exists. Furthermore, negative transfer on the A-C list appears to be associated with, and may be caused by, extinction of the first-list S-R associations. Finally, it seems apparent that retroactive inhibition (RI) in recall which occurs immediately after interpolated learning, will be primarily determined by the amount of extinction occurring during the learning of the second list.

It is an accepted fact that in conjunction with the formation of S-R associations, R-S or backward associations may also develop (e.g., Richardson, 1960). In transfer studies, conflicting backward associations may be responsible for negative transfer in certain paradigms, particularly in the A-B, C-B paradigm (stimuli different, responses identical). It can be seen that if a backward association develops in learning A-B, and if, similarly, a backward association is to develop in learning C-B, the backward association of List 1 (B-A) will interfere with the learning of the backward association in List 2 (B-C). In effect, the backward associations in the A-B, C-B paradigm form an A-B, A-C paradigm. If the R-S associations between two lists form an A-B, A-C paradigm, it seems quite possible that the R-S (B-A) associations of List 1 may also be extinguished in the same fashion as the S-R (A-B) associations. Barnes (1960) has inferred this extinction and she accounts for certain transfer phenomena by it. More particularly, Barnes inferred three different types or classes of associations which are subject to extinction whenever they complete an A-B, A-C paradigm between two lists: (*a*) forward, (*b*) backward, and (*c*) contextual associations. The present study provides a more direct test of the effect of extinction of R-S associations; in effect, it will be a study of RI of backward associations. Four paradigms will be used; if List 1 learning is symbolized as A-B, then the four paradigms may be distinguished in terms of List 2 symbols and their relation to List 1. Expectations concerning the RI of List 1 R-S associations for the four paradigms may now be stated.

A-B, A-C paradigm.—In learning the A-C list, no A-B, A-C paradigm is set up for the backward association, B-A; hence, no extinction of B-A is to be expected. This prediction implies complete independence between the forward and backward associations. However, to the extent that these two associations are nonindependent, RI of the B-A association will be affected by the extinction of the A-B association, since an A-B, A-C relation is obviously present for the forward association. The A-C paradigm offers an assessment of this independence since the only possible

[1] This work was done under Contract Nonr-1228 (15), Project 154-057 between Northwestern University and the Office of Naval Research.

source of extinction of the B-A association lies in the extinction of the A-B association.

A-B, A-Br paradigm.—In this paradigm List 2 contains the same stimuli and responses as List 1, but the stimulus and response terms are paired differently than in List 1. In learning the A-Br list an extinctive relationship (A-B, A-C) is formed for the backward association, i.e., B-A, B-Ar, so that RI of the B-A association is to be expected in this paradigm. In addition, further RI will be produced to the extent that the A-B, A-C relationship for the forward association directly affects the backward association. Since both the A-C and A-Br paradigms contain an extinctive relationship for the forward association while a similar relationship for the backward association is present in the A-Br paradigm only, a comparison between these two paradigms will provide a direct test of the hypothesis that extinction of the backward association can occur.

A-B, C-D paradigm.—This paradigm specifies the learning of two unrelated lists and may provide for a third type of association which is subject to extinction. Although an A-B, A-C relationship exists neither for the backward association nor the forward association for item pairs, there is the possibility of an extinctive relationship being set up between contextual cues and the stimuli of the two lists. In the learning of a list of paired associates, the responses must first be acquired before they enter into an associative connection (Underwood & Schulz, 1960). In the same manner, when the backward association is measured, the stimulus term must necessarily have been acquired as a response if it is evoked. Although it is not known to what stimulus either the response term or the stimulus term are attached when learned in this sense, it can be assumed to be some aspect of the experimental context, e.g., the memory drum, the experimental room, etc. If the association between contextual cues and the stimuli of List 1 are represented as A-B, an A-B, A-C relationship is formed during the learning of List 2 since a different set of stimuli (C) must now be attached to the same contextual cues (A). It is in this manner that the contextual association is thought to provide for a potential class of extinction-susceptible associations. At any rate significant RI of the backward association for this paradigm can be taken as support for the existence of this third class of associations.

A-B, C-B paradigm.—In this paradigm two types of associations which can form an A-B, A-C association may be identified and predicted to contribute to the RI of the B-A association, namely, the backward association and the contextual association. The potential A-B, A-C relationships for these two classes of association already have been elaborated above. As with the first two paradigms listed, the C-D and C-B paradigms allow for a direct test of the extinction of the backward association; that is, the contextual association as a source of extinction is held in common for the two paradigms so that greater RI in the C-B paradigm will indicate direct extinction of the R-S association.

METHOD

Design.—The design was essentially identical to that employed by Barnes (1960). Since an RI paradigm was employed, all *Ss* received the same List 1; the groups were differentiated only by the construction of List 2. The four transfer paradigms of A-B, C-D; A-B, C-B; A-B, A-C; and A-B, A-Br were represented. A control group (Group C) received the single A-B list and was used to measure the amount of stimulus recall possible following an interval equal to the time that List 2 was presented to the transfer groups.

Lists.—The same materials constructed by Barnes were used in the present study. Each list consisted of eight nonsense syllable-adjective pairs. The stimuli for all lists consisted of nonsense syllables of 53% to 73% Glaze association value with intralist similarity and, where applicable, interlist similarity at a minimum. Responses were common adjectives having minimal intralist and interlist meaningful similarity. Four orders of presentation were constructed for each list to minimize serial learning of the responses; each order was used as a starting order an equal number of times for each condition.

Subjects.—Twenty *Ss* were placed in each of the five groups. To assure random assignment of *Ss*, the five experimental conditions were randomized 20 times so that in each of the 20 blocks of 5 *Ss*, each condition occurred once. The *Ss* were assigned to the experimental conditions in the order of appearance in the laboratory. The *Ss* were students of undergraduate psychology courses for whom serving in experiments was a class requirement. Most *Ss* had previously served in other paired-associate experiments; however, none had experienced an RI paradigm nor had been given a procedure involving stimulus recall.

Procedure.—The paired associates were presented on a memory drum by the anticipation method at a 2:2 sec. rate. A 4-sec.

TABLE 1

MEAN R-S RECALL

Groups	Mean	σ_m
A-Br	3.40	.34
A-C	5.05	.43
C-B	1.85	.42
C-D	4.60	.45
Control	5.80	.44

intertrial interval was employed. List 1 learning was taken to a criterion of one perfect recitation. For all groups except the control, a 1-min. rest was given during which time List 2 was positioned on the memory drum. List 2 was administered for 15 anticipation trials. Group C, after reaching List 1 criterion, was given a corresponding 1-min. rest and then instructions for the interpolated "rest" activity (the pyramid puzzle). The pyramid puzzle was continued for 9¼ min., an interval equivalent to the time spent by the transfer groups on List 2. Following the completion of either List 2 learning or the interpolated rest activity, a 2-min. free-recall test was presented in which *S* wrote down the List 1 stimuli opposite the appropriate List 1 responses listed on a dittoed sheet. The ordering of the responses on the sheet corresponded to the ordering that *S* would have received had List 1 learning been continued beyond criterion. An attempt was made to insure that all *S*s understood what was required on the recall test prior to the commencement of the test. No *S* failed to reach List 1 criterion, while 1 *S* was dropped for failure to follow anticipation instructions, and 1 *S* failed to follow recall-test instructions.

RESULTS

The mean number of trials to reach criterion on List 1 for the five groups ranged from 10.55 to 15.10. An analysis of these means, however, failed to reach significance ($F = 1.68$, $df = 4/95$, $P > .05$). Although not of primary interest in the present experiment, the performance of the transfer groups on List 2 revealed Group C-D to be superior to all groups throughout the 15 acquisition trials of List 2, and Group A-Br to be generally inferior. These results are quite in line with those of previous investigators (Barnes, 1960; Twedt & Underwood, 1959).

In measuring for the presence of backward or R-S associations, *S* was given the response terms of the A-B list and was asked to supply the stimulus terms paired with them during original learning. Only if a stimulus term was placed with its appropriate response term was a backward association said to have been demonstrated. The mean number of R-S associations for each group is shown in Table 1. The score for Group C, which is approximately 72.5% of the maximum possible (8.00), represents R-S recall without any apparent possibility of extinction of R-S associations. Inspection of Table 1 shows all transfer groups to have suffered some degree of RI when compared to Group C. An analysis of variance of the recall scores produced highly significant variation ($F = 13.70$, $df = 4/95$, $P < .001$). Comparison of means were made using the within mean sum of squares to estimate the standard error ($\sigma_{diff} = .59$); only the differences between Group A-C and Groups C-D and C did not prove to be reliable.

A problem of method arises in the interpretation of certain of the differences between means. An unavoidable procedural difference between Groups A-Br and A-C and the other three groups is the fact that the former groups received 15 more exposures to the A stimulus terms than did the latter groups. Increased exposure to the A terms during List 2 learning may make these terms more available to *S* at recall. This does not mean that the B-A associations will necessarily be stronger, but only that the A terms might be more available as responses at recall. This, in turn, might increase the number of "hits" if *S* guessed. There is, thus, the distinct possibility that, relative to

Groups C-B and C-D, the amount of RI for Groups A-Br and A-C is underestimated.

The results for the four transfer paradigms may now be assessed in terms of the three classes of associations thought to be susceptible to extinction whenever an A-B, A-C association is formed. The fact that significant RI was not obtained in Group A-C ($t = 1.27, df = 95, P > .05$) suggests that extinction of the A-B association does not result in the extinction of the corresponding B-A association as indexed by subsequent R-S recall. It is possible, however, that the methodological problem discussed above would result in an overestimation of the amount of recall for Group A-C through an increase in the number of stimulus terms available to *S*. Without any data concerning the extent to which the recall scores are inflated by guessing, it can only be tentatively concluded that extinction of the S-R and R-S associations represent two independent processes.

A comparison of the recall for Groups A-Br and A-C will not be biased due to differential stimulus-term exposure because both groups had the same number of exposures. Since both groups contain the S-R extinction factor in common, the significant difference in R-S recall for these two groups ($t = 2.80$, $df = 95$, $P < .01$) may be taken as one indication of an extinction of R-S associations when these associations form an A-B, A-C paradigm.

The final critical comparisons are between Groups C-D and C-B. Since Group C-D contains only the contextual extinction factor and Group C-B both the contextual and R-S extinction factors, the greater RI for Group C-B ($t = 4.66, df = 95, P < .01$) indicates again that the R-S association can be extinguished. The fact that the contextual associations are involved in the recall of Groups C-B and C-D is shown by the comparison between Groups C-D and C ($t = 2.03$, $df = 95$, $.01 < P < .05$).

Discussion

This study was designed to study RI of backward (R-S) associations. The evidence presented makes it clear that when R-S associations between two successively learned lists form an A-B, A-C paradigm, RI of List 1 R-S associations may occur. Indeed, in the A-B, C-B paradigm, the magnitude of the RI was great; a difference of about four items was apparent between the control condition and recall of List 1 R-S associations with this paradigm. In line with certain previous interpretations of RI of S-R associations (Barnes & Underwood, 1959), the RI of R-S associations in the present experiment has been interpreted as being due to an extinction or unlearning of the R-S associations of the first list during the learning of the second. Certain issues related to such an interpretation require further discussion.

The fact that the A-Br paradigm resulted in a significant decrement in recall of R-S associations strongly implies an extinction process of specific R-S associations in the first list. Indeed, the amount of decrement observed is probably underestimated in the present experiment. As noted earlier, stimulus terms should be more available in this paradigm than in the control condition because of the greater number of trials they were presented. This could result in more correct R-S pairings based on guessing alone. That guessing might have occurred is suggested by the fact that in 22 instances *S*s gave a correct stimulus term but paired it with the wrong response term, and in only 1 of these 22 cases was the pairing appropriate to the association learned in List 2. Of course, if guessing did occur it may have been heightened because R-S associations of the first list had indeed been extinguished; *S*s knew the stimulus terms but extinction of the specific R-S associations

left them with no appropriate pairing tendencies. Such an interpretation might be supported by the fact that in the A-C paradigm only 9 stimulus terms were paired incorrectly and in this paradigm frequency of stimulus-term exposure was equal to that of the A-Br paradigm, but no extinction of R-S associations should have occurred.

In addition to extinction of R-S associations between pairs of items in the two lists, it has been suggested that another source of extinction may contribute to the observed recall decrement, namely, extinction of contextual associations. The extinction of such associations is said to be the sole source of decrement in the C-D paradigm. For this paradigm it is said that in learning A-B, the stimulus term A is originally associated with contextual cues, but that during C-D learning the associations between the A terms and the contextual cues are extinguished and are replaced by associations between the C terms and contextual cues. It should be noted that this process alone will not account for the decrement in recall of the B-A association for this association per se has not been extinguished. It is clear, therefore, that further assumptions are needed in order to mediate the findings if extinction of contextual associations is to be retained as a basic notion. There are several alternative assumptions which may be added and which would satisfactorily mediate the findings. However, it seems most apparent that the critical need at the present time is not further assumptions; rather, what is needed is experimental simulation of contextual cues (hence associations) in order to bring them under laboratory control and thus determine whether or not the basic notion (extinction of contextual associations) is sound.

Finally, it should be pointed out that just as in the case of presumed extinction of S-R associations (Barnes & Underwood, 1959), the reason why all R-S associations are not extinguished is not known.

Summary

This experiment investigated RI of backward or R-S associations. Various interlist relationships were studied on the premise that first-list R-S associations, like S-R associations, would be subject to extinction if the R-S associations for the two lists formed an A-B, A-C interference paradigm. All *Ss* learned the same first list of paired associates which had nonsense syllables as stimulus terms and adjectives as response terms. Four different groups (20 *Ss* each) then learned a second syllable-adjective list, the groups being differentiated by the nature of the interlist paradigm formed, these being: C-D, C-B, A-C, and A-Br. Following acquisition of the second list, *S* was allowed 2 min. to write down the stimulus terms from the first list opposite the appropriate first-list response terms. A control group was given no second list.

All groups showed significant RI in R-S recall except the group learning the lists forming the A-B, A-C paradigm. No source of extinction of R-S associations for this paradigm is apparent and extinction of the S-R association does not affect the R-S association. Maximum RI was shown for A-B, C-B; this paradigm may involve extinction of specific R-S associations as well as contextual associations. The A-Br paradigm, involving only extinction of specific R-S associations, produced more RI than did the A-B, C-D paradigm; this latter paradigm is assumed to involve a decrement resulting only from extinction of contextual associations.

References

Barnes, J. M. "Fate" revisited. Unpublished doctoral dissertation, Northwestern University, 1960.

Barnes, J. M., & Underwood, B. J. "Fate" of first-list associations in transfer theory. *J. exp. Psychol.*, 1959, **58,** 97–105.

Briggs, G. E. Acquisition, extinction, and recovery functions in retroactive inhibition. *J. exp. Psychol.*, 1954, **47,** 285–293.

Richardson, J. Comparison of S-R and R-S learning of paired-associates. *Psychol. Rep.*, 1960, **7,** 225–228.

Twedt, H. M., & Underwood, B. J. Mixed vs. unmixed lists in transfer studies. *J. exp. Psychol.*, 1959, **58,** 111–116.

Underwood, B. J., & Schulz, R. W. *Meaningfulness and verbal learning.* Chicago: Lippincott, 1960.

(Received September 15, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 405-409

CUE SELECTION IN PAIRED-ASSOCIATE LEARNING

BENTON J. UNDERWOOD, MARGARET HAM, AND BRUCE EKSTRAND

Northwestern University

Consider a paired-associate list in which the stimulus term for each response consists of two distinct components, A and B. Both components are consistently present on each learning trial. Assuming that learning occurs in this situation, there are many possible interpretations which may be given as to what the effective stimulus or cue is for each response. It might be said that the effective cue is a configuration formed by A and B. Or, it might be said that each component is independently a cue for the response; that one or the other components is the effective cue, but not both, and so on.

The present study is predicated on the notion that when a complex stimulus is presented to *S* a selection process may occur so that the effective cue for the response is some component of the complex stimulus that is actually presented. Thus, the assumption is that there may be a discrepancy between the nominal stimulus (the stimulus actually presented *S*) and the functional stimulus (the component of the nominal stimulus which becomes the effective cue for response elicitation). That such discrepancies may exist is suggested by the reports of *S*s that they have used only a single letter of a three-letter stimulus as the effective cue (Underwood & Schulz, 1960). Such discrepancies might also be inferred from the so-called context experiments (e.g., Weiss & Margolius, 1954) in which the removal of a component of a compound nominal stimulus produces a decrement in recall, although such studies offer other interpretative possibilities, e.g., the functional stimulus is a configuration and the removal of any component reduces the effective associative strength.

If cue selection occurs—if only a part of a compound stimulus becomes the functional stimulus—certain variables should influence the selection. The hypothesis tested in the present experiment is that given two components of different classes as the nominal stimulus, the more meaningful component will become the functional stimulus. This hypothesis seems very close to the notion of differences in discriminability as a variable determining stimulus selection, a notion suggested by Sundland and Wickens (1962), and for which some experimental support was obtained.

The particular predictions for the present study may now be specified. Two lists for original learning were constructed. The stimulus compound for one list consisted of colors and low-meaningful trigrams; for the other list the compound consisted of colors and common three-letter words. For the first list it was assumed that the colors were more meaningful than the trigrams; therefore, the functional stimuli should be the colors. In the case of the word-color compound it was assumed that the words were more meaningful than the colors, hence, the functional stimuli should be the words. (It would have been more precise to have used two sets of verbal units of known meaningfulness for the compounds, but the use of colors was recommended by the desire to keep the experiment continuous with the context experiments.) Given the

above assumptions, it was predicted that following the learning of the list with the trigram-color compounds, very little decrement would be observed if the trigrams were removed on a transfer test, but that a great loss would appear if the colors were removed from the compound. Contrariwise, in the case of the word-color compounds, removal of the words on a transfer test would result in a large loss but removal of the colors would have little effect on transfer performance.

METHOD

The general procedure required that half the *Ss* learn an original list with trigram-color compound stimuli, and half learn a list with word-color stimuli. To test for each component separately on a transfer list required four groups. However, to determine precisely the amount of decrement occurring required two control groups for which the stimuli on the transfer test were the same as on original learning. Thus, the design called for six groups. Let W stand for words, T for trigrams, and C for colors, and the symbols before a hyphen designate the stimulus during original learning, those after the hyphen the stimulus on the transfer test. The six groups are, therefore: WC-WC and TC-TC (the two control groups); WC-C and TC-C (only color stimuli on transfer test); WC-W and TC-T (the verbal units appear alone as the stimuli on the transfer test).

Lists.—The materials for the lists are shown in Table 1. Each list consisted of seven pairs. The three-letter words all have AA ratings in the Thorndike-Lorge (1944) list except the word GAS, which is rated A. The trigrams have quite low associative connections between letters as based on the Underwood-Schulz (1960) tables. It should also be noted that the trigrams have relatively high formal similarity as indexed by repeated letters. The purpose of this was to minimize the possibility that a single letter (such as the first letter) might become the functional stimulus. The frequency of repeated letters in the word list is about the same as for the trigrams and the repetitions are in the same positions. Both lists have the same initial letters.

TABLE 1

STIMULUS COMPONENTS USED IN THE LISTS

Words	Trigrams	Colors
GAS	GWS	Red
DAY	DWK	Brown
NEW	NXQ	Yellow
DIE	DHX	Blue
BAD	BWD	Orange
GOT	GVS	Black
BED	BXD	Green

TABLE 2

MEAN NUMBER OF TRIALS TO CRITERION ON ORIGINAL LEARNING AND MEAN NUMBER OF ITEMS LOST ON FIRST TRANSFER TRIAL

Cond.	Original Learning		Transfer	
	Mean	σ_m	Mean	σ_m
WC-WC	8.80	.97	—	—
WC-W	8.00	1.18	1.20	.44
WC-C	9.20	1.14	2.50	.47
TC-TC	11.55	.95	—	—
TC-T	10.00	1.20	2.85	.36
TC-C	10.00	1.21	.05	.37

The color components were made of construction paper and pasted on the vellum tape. Rectangular frames of color completely surrounded the verbal unit, the width of the frame being approximately $\frac{1}{4}$ in. When the color was the only component on the transfer test the frame appeared exactly as it had during original learning, the verbal unit having been removed. The particular colors paired with particular verbal units appear in the same row in Table 1. The response terms were the single digit numbers 2 through 8.

On both learning and transfer trials the rate of presentation was 2 sec. for the stimulus alone, and 2 sec. for the stimulus and response together. Four different orders of the pairs were used. Anticipation learning was used throughout.

Procedure.—The original learning was carried until the *S* achieved one perfect recitation of the list. The transfer tests were carried for 10 trials with *S* instructed to give as many correct responses as possible on the first transfer trial. The usual paired-associate instructions were given prior to original learning. In addition, *S* was told that both the word (or trigram) and color would appear consistently together from trial to trial. The intent of these instructions was to inform *S* of the nature of the stimulus compound with-

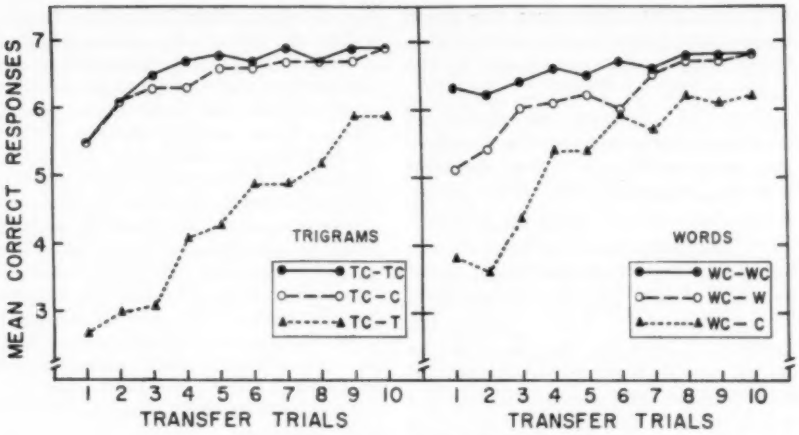

FIG. 1. Acquisition curves on the 10 transfer trials.

out, at the same time, biasing him toward "using" one or the other components. Prior to the transfer trials *S* again was fully informed as to the nature of the stimulus which would be present on these trials. Approximately 45 sec. elapsed between original learning and the first transfer trial.

Each of the six groups contained 20 *Ss*. Twenty blocks were made up such that each condition occurred once within each block, with the order of the six conditions within a block being randomly determined. The *Ss* were then assigned to the schedule in terms of their appearance at the laboratory. No *S* was lost for failure to learn.

RESULTS

The mean numbers of trials to attain the criterion on original learning are shown in the left portion of Table 2. Differences among the three WC groups and among the three TC groups represent random variation, the F being less than 1 in each case. For the 60 WC *Ss* the mean is $8.67 \pm .63$, and for the 60 TC *Ss*, $10.52 \pm .65$. The difference ($1.85 \pm .90$) gives a t of 2.06, which is just past the 5% significance level.

The mean performance on each transfer trial is shown in Fig. 1. The left-hand section refers to the TC groups, the right-hand section to the WC groups. For the TC groups it can be seen that when the colors alone were used as stimuli, transfer was virtually complete; the performance of this group is only slightly below the control (TC-TC), thus indicating that the colors alone were completely effective functional stimuli. If *only* the color component was the functional stimulus, performance of Group TC-T should start at zero. It does not. For at least some *Ss* the trigrams were also functional stimuli for at least a few responses.

The right-hand section of Fig. 1 shows that for the WC groups neither the words nor the colors developed complete effectiveness. The colors are less effective than the words, but the words alone show some loss as compared with the control.

The clearest inferences concerning the functional stimuli in original learning can be made from the performance on the first transfer trial. Since Group TC-TC showed a larger criterion drop than did Group WC-WC, loss scores have been calculated. To do this, each *S*'s score on the first

trial was subtracted from the mean score of the appropriate control group. These loss scores are shown in the right section of Table 2. A 2 × 2 analysis of variance was performed on these scores, using as one classification variable TC and WC as identified in original learning, and as the other, colors and verbal units on transfer. Only the interaction F was significant, being 25.31; with 1 and 76 *df*, the F needed for the 1% significance level is approximately 7.00. Thus, the predictions that for the TC lists the colors would become the functional stimuli and that for the WC lists the words would become the functional stimuli, are given some support.

Discussion

The transfer tests for the TC compounds showed that color alone was a completely effective functional stimulus. This fact precludes an interpretation of the functional stimulus as being a configuration. However, it was noted that the trigram stimuli were not completely ineffective on the first transfer trial. There are at least two possible interpretations of this finding. First, it may mean that some trigrams, quite independently of the color component, become associated directly with the response term. Secondly, it may mean that some associative connection may have developed between the color and trigram components of the stimulus compounds. Thus, when the trigram is presented, the correct anticipation is mediated by associations running from the trigram to the color to the response term. The present data do not allow a choice between these two alternative interpretations.

The transfer data following learning of the WC lists raise three interpretative problems. First, it was noted that transfer was greater when the words became stimuli than when the colors became stimuli. Two circumstances could lead to this finding. (*a*) Most *S*s in original learning used words as stimuli for *all* associations but a few *S*s used colors as stimuli for all associations. (*b*) All *S*s used words as stimuli for *most* associations during original learning but used color as stimuli for a few associations. Given a large number of *S*s a choice between these two alternatives could be made by examining the distributions of scores on the first transfer trial. If the first of the two possible explanations is appropriate, the distributions should be bimodal when words alone or when colors alone are stimuli on the first transfer trial. If the second alternative is appropriate, each distribution should be continuous. Actually, bimodality is suggested in the present distributions but with only 20 cases in each this may be quite fortuitous.

The second interpretative problem is the same as that posed for the TC compounds where the data show that for some associations for some *S*s both the trigram and the color elicited the response. Such dual functionality may also be deduced from the data for the WC compounds. On the first transfer trial a mean of approximately 5.0 correct responses occurred when the words were presented alone and 3.8 when the colors were presented alone. These two values sum to 8.8, which is appreciably higher than the mean of 6.3 shown by the control *S*s on the first transfer trial. Clearly, dual functionality of the two components obtained for at least some *S*s. However, just as in the case of the TC compounds, this apparent duality may result from direct associations between the components and the response term or it may result from mediation between the stimulus components.

The third interpretative problem presented by the results of the WC compounds is the fact that the words showed greater transfer than did the colors. As stated in the procedure section, it was believed that the words would be more meaningful than the colors. We have no independent evidence for this and it may not be valid. The *S*s, being much more practiced in dealing with word stimuli than with patches of color as stimuli, may be biased toward the selection of the

verbal stimuli. The data are quite in harmony with such a notion. However, a somewhat different approach may be taken to the problem. An empirical test can be made to determine which stimulus compound leads to most rapid learning when this learning is not preceded by learning in which the compound is present. To determine this three new groups of 15 *S*s each were run. One group learned the trigram-number pairs, a second the word-number pairs, and the third the color-number pairs. The mean total correct responses in 10 trials were 38.80 ± 3.61, 51.20 ± 3.52, and 50.40 ± 2.19, respectively. While the F is significant far beyond the 1% level it is clear that most of the variance is produced by the trigram-number pairs. The words and colors do not differ appreciably in their effectiveness as stimuli. The small difference in favor of the words occurred primarily on the first three trials. Thus, it seems quite reasonable to conclude that when *S* is given a compound consisting of common words and colors, he is likely to select the words as functional stimuli, not necessarily because they are more meaningful, but because he is more accustomed to dealing with such stimuli.

Finally, it may be stated that the results of the present experiment, taken in conjunction with the study by Sundland and Wickens (1962), would seem to indicate that it may be more fruitful to view the so-called context experiments as experiments investigating the variables determining cue selection.

Summary

This experiment was based on the assumption that when *S* is presented a compound stimulus in a verbal-learning experiment, cue selection may occur. Word-color or trigram-color compound stimuli were used in learning original paired-associate lists with numbers as responses, followed by a transfer test in which one or the other components alone was presented as the stimulus. Control groups were also used, these groups being given further trials with the original compound stimuli.

The results show:

1. For the trigram-color compounds, color was a completely effective stimulus on the transfer test. The trigrams, however, also produced a small positive transfer effect. The selection of the color component as the primary functional stimulus was assumed to be due to its higher meaningfulness.

2. For the word-color compounds, transfer was higher when the words appeared alone than when the colors appeared alone. This may be due to a bias *S*s have toward dealing with verbal material (as compared with the color patches used) rather than to higher meaningfulness of the words.

It was concluded that experiments dealing with the effects of context changes on retention may be viewed as representing cases of cue selection.

References

Sundland, D. M., & Wickens, D. D. Context factors in paired-associate learning and recall. *J. exp. Psychol.*, 1962, **63**, 302–306.

Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words.* New York: Teachers College, Columbia University, 1944.

Underwood, B. J., & Schulz, R. W. *Meaningfulness and verbal learning.* Chicago: Lippincott, 1960.

Weiss, W., & Margolius, G. The effect of context stimuli on learning and retention. *J. exp. Psychol.*, 1954, **48**, 318–322.

(Received September 15, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 410–414

CONSUMMATORY AND INSTRUMENTAL RESPONDING AS FUNCTIONS OF DEPRIVATION[1]

GEORGE COLLIER

University of Missouri

The present study investigates the conjoint effect of deprivation on instrumental and consummatory responses.

METHOD

Apparatus.—Eight Skinner boxes delivering liquid reinforcement were used (Collier & Myers, 1961). Electronic switches sensed each lick made in the magazine cup. Bar presses (BP), reinforcements, and licks were recorded by an Esterline-Angus recorder.

Subjects.—The 12 *S*s were naive rats, 120 days old, of the Sprague-Dawley strain (Holtzman Company) maintained on Purina lab chow and tap water.

Procedure.—Two groups of six rats were formed. The first (H) was fed for 1 hr. after running. The second (L) had its food removed 3 hr. before running and replaced immediately after running. From Day 14 to Day 17 Group H received 15 mg. of food per day. From Day 17 to the finish they received 10 gm. per day. These changes were required to keep *S*s' weights relatively constant.

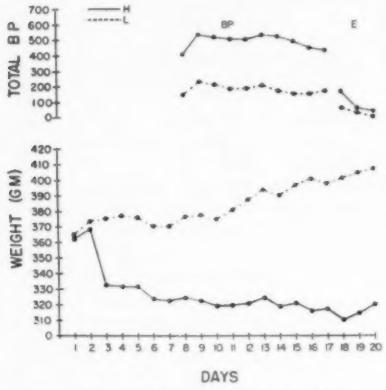

FIG. 1. Total daily bar presses and weight of *S*s as a function of deprivation.

Each *S* received 5 days of table training, handling, and deprivation accommodation followed by 4 days of magazine training, 10 days of reinforced BP on a 1-min. fixed interval (FI), and 3 days of extinction. During reinforced bar pressing all *S*s were given .3 ml. of 16% sucrose solution per reinforcement. Thus, since by the fifth day of bar pressing all reinforcements possible were presented and all those presented were taken, each *S* received 15.6 ml. of solution per day in the 52-min. test period.

All *S*s were weighed every day immediately before running.

RESULTS

Reinforced bar pressing.—The lower panel of Fig. 1 presents the weights of the two deprivation groups over the course of the experiment. Group H ranged from approximately 86% of the body weight of Group L at the beginning of the study to approximately 80% at the end. The weight of Group L increased significantly across sessions. The upper panel of Fig. 1 presents the total number of BP in 52 min. for each of the deprivation groups over the course of the experiment. It is clear that Group H was superior to Group L under all circumstances.

An analysis of variance of the number of BP and of licks on Day 17 shows that for bar pressing the main effects of Deprivation ($F = 22.6$, $df = 1/10$, $P < .01$) and of Minutes ($F = 2.43$, $df = 51/510$, $P < .01$) were significant, while for licks the main effect of Minutes only ($F = 1.46$, $df = 51/510$, $P < .05$) was significant. For neither measure did the interaction of Deprivation × Minutes

[1] This investigation was supported in part by Research Grant M-3328, from the National Institute of Mental Health, Bethesda, Maryland.

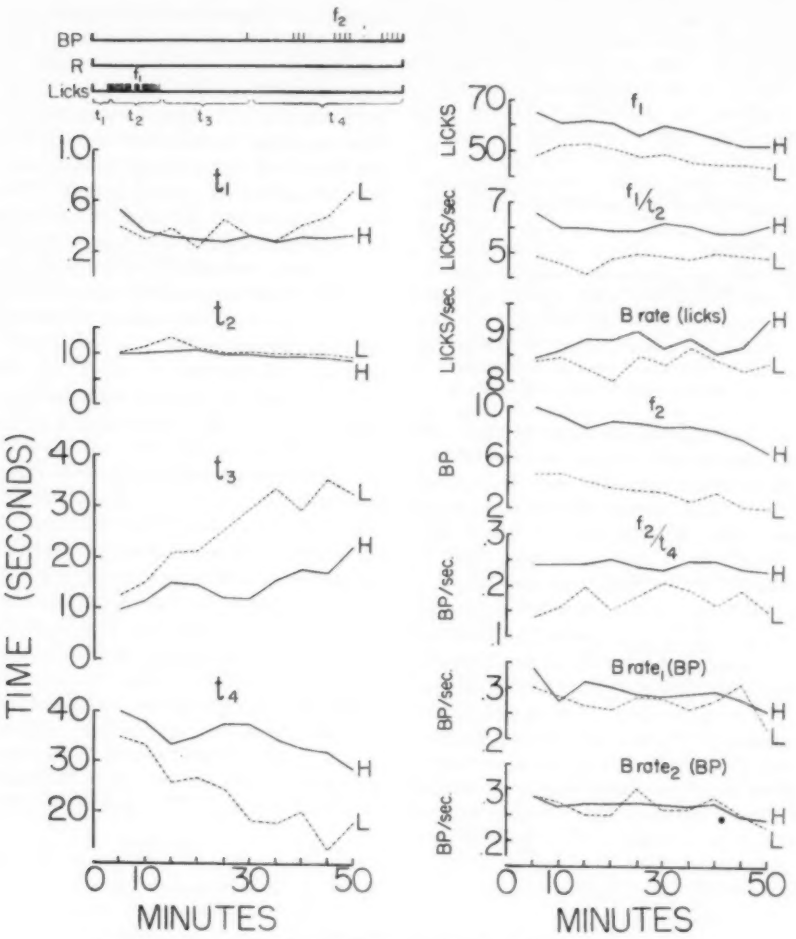

FIG. 2. Latency, duration, rate, and burst rate of licking, and bar pressing as functions of deprivation and satiation.

approach significance. The total numbers of BP per session were 439.7 and 174.0 for Groups H and L, respectively, while the total numbers of licks were 2763 and 2400, respectively.

In order to assess the locus of the rate differences and changes observed, the pattern of responding in each minute of the FI schedule, i.e., between reinforcements, was examined. Figure 2 (top left) presents a schematic of 1 min. (between reinforcements) of a typical tape from the Esterline-Angus records: t_1 is the time from the reinforcement to the first lick, t_2 is the duration of licking, t_3 is the time from the last lick to the first BP, t_4 the time from the first BP

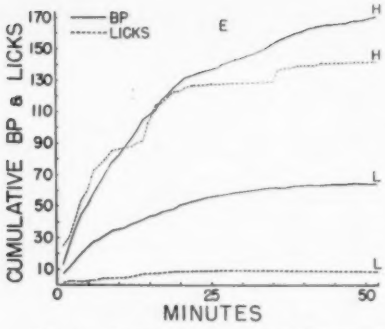

FIG. 3. Cumulative number of bar presses and licks over the course of extinction as a function of deprivation and satiation.

to the next reinforcement; f_1 is the number of licks, f_2 the number of BP. The average rate of licking (f_1/t_2) and of BP (f_2/t_4) as well as the burst rates were also determined. A burst was defined as a sustained sequence of licks or BP without pause, where a pause is an interval between responses twice that of any preceding interval (as determined by visual inspection). Because of the difficulty of identifying a BP burst, an additional measure, the reciprocal of the time between the last two responses before reinforcement (B-rate$_2$), was calculated as a check. Figure 2 presents the values averaged over 5-min. blocks. Analyses of variance on each of the measures were performed for deprivation and minutes of the session.

Summarizing these results, no significant effect of Deprivation on the average (H = 6.0/sec, L = 4.8/sec) or Momentary (H = 8.8/sec, L = 8.4/sec) Rate of Licking was found; however, there was a suggestion that there might be a difference in average rate as a function of a greater number or duration of pauses between bursts of licks at a constant rate ($F = 4.35$, $df = 1/10$, $P < .10$). The initial latencies of the first licks were significantly longer at the beginning and tended to increase significantly toward the end of the session. The average duration (10.1 sec.) of licking was unaffected by the course of satiation. That some licking beyond that required occurred is indicated by the fact that the average volume per lick was .0056 ml. per lick and .0065 ml. per lick for Groups H and L, respectively, where the typical value in the unrestricted access situation is usually of the order .0080 ml. per lick. The average rate declined significantly over a session while the burst rate appeared to increase.

The latency of the first BP following reinforcement (t_3) was significantly shorter in Group H (15 vs. 26 sec.) and increased significantly over the session for both groups. The duration (t_4) of BP was, therefore, necessarily significantly longer (35 vs. 23 sec.) in Group H and significantly decreased over the session for both groups. For neither measure did the Deprivation × Minutes interaction approach significance, suggesting that the rate of decline in responding was independent of the degree of deprivation. No evidence was found for a difference between within-cluster (burst) rates on either measure (B-rate$_1$: H = .29/sec, L = .27/sec; B-rate$_2$: H = .27/sec, L = .24/sec). However, both the number (f_2) and the average rate (f_2/t_4) were significantly greater for Group H (8.4 vs. 3.5 and .24/sec vs. .17/sec, respectively). The average rate did not decline over the session while the frequency of responding did. The difference between the average rate of responding and the momentary rate reflects either the burst duration, interburst interval, or both. The decline in frequency across the session appears to be solely due to the decreasing time available for BP because of the increasing latency of the first BP.

Extinction.—The cumulative form of the extinction curve is given in Fig. 3. Both the total number of BP and total number of licks differed significantly as a function of deprivation. The rate of decline in bar pressing was significantly faster for Group L as was the interaction of minutes and deprivation. No analysis of the effect of time and its interaction with deprivation was performed for licks since Group L stopped licking within the first few minutes.

Two things were clear in the extinction data. First, the rate at which consummatory responding as well as bar pressing declined depended strongly on deprivation, stopping with a few licks of the empty magazine at low deprivations. Second, bar pressing fell off at a much slower rate than licking, substantial amounts of bar pressing occurring after the licking response to the magazine had stopped.

Discussion

When the distribution of responses between reinforcements is examined it is clear that the momentary rate of responding for both the consummatory and instrumental response is relatively insensitive to either deprivation or the number of reinforcements consumed. The difference in average rate as a function of deprivation is the result of differences in the duration, or interval between bursts, or both, and the differences in overall rate of BP as a function of deprivation result from both the latency of the first response following reinforcement and the differences in average rate. The decline in the overall rate of BP as a function of the number of reinforcements consumed appears to be solely a function of the latency of the first response following reinforcement and the rate of decline is independent of the degree of deprivation. This distribution of responses has been reported previously for the consummatory response when deprivation and amount consumed are examined (e.g., Davis & Keehn, 1959; Stellar & Hill, 1952) and for the instrumental response on a fixed-ratio schedule (Sidman & Stebbins, 1954). The relatively small difference between average rates of licking, the fixed duration of licking, and the apparent lack of difference in the latency of the first lick probably result from the schedule used in which a fixed, small volume is presented which can be consumed in a single or a few brief bursts of licking. That is, the duration required is within the range of durations of bursts of licking. The failure of the latency of the first lick to vary as a function of the amount consumed (see Stellar & Hill, 1952) is probably the result of the fact that the consummatory response in the Skinner box situation is under the control of a precise discriminative stimulus, namely the operation of the magazine.

Ingestion ceases before nutrition has been accomplished. This fact suggests that satiation and deprivation are two distinct, independent processes (Collier & Myers, 1961). The results of the present study provide further evidence for this distinction. They suggest that satiation, which varies as a function of the momentary postingestive load, evinces its major effect on the latency of the response measured, while deprivation affects both the latency and burst density. The present data further suggest the possibility that the parameters of the consummatory response are insensitive to such variables as deprivation and satiation, and thus it is the ancillary food approaching and food producing responses which determine the rate, pattern, and amount of consumption.

Summary

Twelve albino rats were divided into two deprivation groups, high and low, and were run in Skinner boxes for a .3-ml., 16% sucrose reinforcement. Analysis of the grain of the consummatory (licks) and instrumental (bar press) responses led to the conclusion that for the consummatory response neither the

latency, duration, average, nor momentary rate of responding varied significantly as a function of deprivation. Latency showed a slight tendency to decrease for high deprivation and increase for low deprivation, average rate to decrease, and momentary rate to increase across a session. For the BP response, on the other hand, the latency and therefore necessarily the duration of pressing were functions of both deprivation and satiation. The average rate proved to be related to deprivation only, while two measures of the momentary rate appeared related to neither deprivation nor satiation. In extinction the rate of occurrence of licking showed a much more rapid decline than did rate of BP.

References

COLLIER, G., & MYERS, L. The loci of reinforcement. *J. exp. Psychol.*, 1961, **61**, 57–66.

DAVIS, J. D., & KEEHN, J. D. Magnitude of reinforcement and consummatory behavior. *Science*, 1959, **130**, 269–271.

SIDMAN, M., & STEBBINS, W. C. Satiation effects under fixed-ratio schedules of reinforcement. *J. comp. physiol. Psychol.*, 1954, **47**, 114–116.

STELLAR, E., & HILL, J. H. The rat's rate of drinking as a function of water deprivation. *J. comp. physiol. Psychol.*, 1952, **45**, 96–102.

(Received October 1, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 415–420

DEPTH PERCEPTION IN ROTATING DOT PATTERNS: EFFECTS OF NUMEROSITY AND PERSPECTIVE[1]

MYRON L. BRAUNSTEIN[2]

University of Michigan

The changing shape of a visual pattern may elicit a three-dimensional perception, even when other cues to depth are so reduced that any stationary view of the pattern appears two-dimensional. Stimuli for the study of this effect have generally been produced by rotating an object or pattern between a light source and a translucent screen, thus displaying shadows of the transforming pattern. This method was employed by Miles (1931), who used a two-bladed electric fan as a shadow caster, and by Metzger (1934), who used a number of cylinders to generate his displays.

The effect was systematically investigated by Wallach and O'Connell (1953), who termed it the "kinetic depth effect," and by Gibson and Gibson (1957). While earlier investigators regarded the depth effects created by such displays as illusions, recent experimenters have recognized that they involve a cue to depth not fully incorporated into the classical set of depth cues. Gibson (1950) has aptly termed this cue "motion perspective."

Green (1959a, 1959b) introduced a method which allows greater variety in the presentation of stimuli and greater control of the motion of parts of a display than the shadow projection method. By means of instructions to a high speed digital computer equipped with a CRT output recorder, a motion picture can be made of a two-dimensional projection of any mathematically specifiable transformation of any figure. Green has studied the effects of numerosity, speed, and axis of rotation on subjective judgments of the extent to which the parts of the display maintain the same relative position, i.e. coherence, for filmed sequences representing projections of points or of line segments rotating in three-dimensional space.

Definition of perspective.—The perspective used in producing a two-dimensional projection of a three-dimensional display may be conveniently defined as the ratio of the distance between the projection point and the most distant $X-Y$ (frontal parallel) plane to the distance between the projection point and the closest $X-Y$ plane, along the Z axis (line of sight). This is equivalent to the ratio of the projection of a distance on the closest $X-Y$ plane to the projection of the same distance on the most distant $X-Y$ plane. Figure 1 illustrates these definitions of perspective for a four-point pattern confined to an imaginary cube.

If the projection point is at infinity and thus equally distant from all elements of the pattern, the perspective ratio is 1, and the projection of the pattern does not change with its location in depth. This method of

[1] This paper is based on a dissertation submitted to the Department of Psychology at the University of Michigan in partial fulfillment of the requirements for the PhD degree. The author gratefully acknowledges the generous assistance of B. F. Green, Jr. of the MIT Lincoln Laboratory in the preparation of the stimuli, and the valuable guidance of W. L. Hays in the planning and performance of this research.

[2] Now at Cornell Aeronautical Laboratory, Incorporated, Buffalo, N. Y.

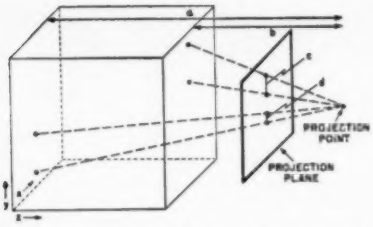

FIG. 1. Projection of points in a cube. (Perspective is defined as a/b, or equivalently, c/d.)

projection is referred to as parallel projection. If the ratio is greater than 1, the projection of the pattern varies with its location in depth, and the projection is referred to as polar.

Motion perspective.—Motion perspective is a complex cue to depth involving several better known aspects of perception. If a polar projection of a three-dimensional pattern rotating about an axis perpendicular to the line of sight is displayed, the rate of motion of the projection of a part of the pattern is an inverse function of its distance from the projection point. This is essentially motion parallax, in the classical sense. Also, in the case of a polar projection, the distance between the projection of two parts of a pattern is an inverse function of the distance of those parts of the pattern from the projection point. This cue is similar to linear perspective, except that it involves temporal as well as spatial variations in the display.

With both polar and parallel projections, changes in the relative distances between the projections of parts of the pattern correspond to changes in orientation of the pattern from a fixed projection point. In the case of parallel projections, this effect is not confounded with motion parallax and linear perspective. Although only very distant objects are seen in parallel projection in direct vision, the present study compares polar and parallel projections in order to separate the more familiar depth cues from the novel aspects of motion perspective.

As the projection point is equivalent to the nodal point of the observer's eye in retinal projections, variations in perspective are normally accompanied by changes in the visual angle subtended by the pattern. In the present study the size of the projection was controlled by placing *O*'s eye at a fixed intermediate distance from the computed location of the three-dimensional pattern, and not at the computed projection points. Perspective is thus "inappropriate" to *O*'s distance from the computed location of the pattern. It is exaggerated by closer projection points and reduced by more distant ones.

METHOD

Stimuli.—The stimuli were produced by Green's (1959a) computer method. Specifically, from two to six points in a cube having sides of two units, centered about the origin, were randomly selected. A projection of the points onto a specified plane perpendicular to the Z axis was computed, using a projection point having a Z coordinate of 2, 4, 16, or 512, and X and Y coordinates of 0. The projection plane was always equidistant from the projection point and the origin so that a distance d on the plane $Z = 0$ was projected into a distance $d/2$, regardless of the distance of the projection point from the origin. Projections were computed at each 4.5° as the hypothetical cube was rotated 360° about the Y axis. There were thus 80 projections for each display. Each projection was plotted on the face of a CRT using an overlapping cluster of four spots to represent a point. These clusters, hereafter referred to as spots, were roughly circular in shape and were approximately .01 unit in diameter. The plots were photographed on 16-mm. film and each sequence of 80 photographs became a stimulus display in the final film.

Forty such displays were prepared, representing 20 treatment combinations (5 levels of numerosity and 4 of perspective), each

produced twice, with different selections of random points. The stimulus films consisted of pairs of these displays. Each pair was composed of an 80-frame display, 28 frames of blank film of the same optical density as the background of the displays, a second 80-frame display, and 108 frames of blank film, again of the same density as the background of the displays. Presented at 16 frames per sec., this would mean a 5-sec. display, a 1.75-sec. pause, another 5-sec. display, and a 6.75-sec. pause, followed immediately by the first display in the next pair.

Apparatus.—The films were projected by a Kodak Royal 16-mm. projector in which a 500-w. bulb was used. The lens had a focal length of 2.5 in. The displays were projected onto a translucent plastic screen, 174⅞ in. from the front of the lens. One unit on the theoretical projection plane became 3.76 in. on the screen. The *O* viewed the displays from the opposite side of the screen. Background and spot luminances were approximately .009 and 1.1 ft-L., respectively. The projector operated at a speed of 16.2 ± .1 frames per sec. The *E* operated the projector and recorded *O*'s responses, which were transmitted to a tape recorder in *E*'s cubicle.

The *O* viewed the displays monocularly through a reduction tunnel. This consisted in part of a cylinder 4⅞ in. in length and 2½ in. in diameter, against which *O* placed his eye. The cylinder was covered on the outside with black tape and lined with black paper on the inside. It was inserted into a circular aperture 2⅜ in. in diameter in a wooden disk 15 in. in diameter and 1½ in. thick. The length of the cylinder inserted into the aperture was ¾ in., leaving 4⅛ in. protruding on the side on which *O* viewed the displays. Black cloth, formed roughly into a cylinder, covered the space between the disk and the aperture in the wall in front of *O*, which was also 15 in. in diameter. The distance from the front of the disk to the wall was 21⅞ in. The *O*'s eye was thus approximately 26 in. from the screen, or 13.8 theoretical units from the origin of the display. The *O* could see neither the borders of the aperture in the wall as he looked into the reduction tunnel, nor the location of the screen as he entered the room. The *O*s were 31 male students enrolled in the elementary psychology course; 20 served in Exp. I and 11 in Exp. II.

Procedure.—Each *O* was tested individually in a 2-hr. session. In Exp. I, *O* was told that he would see pairs of displays, either of which could be seen as occurring in space or on a flat surface in front of him, and was to indicate, by saying "first" or "second," which display in each pair gave the stronger impression of occurring in space. After a practice reel, *O* was shown four reels containing the 190 pairs of the present experiment, as well as 34 additional pairs not included in this report.

The procedure in Exp. II was identical to that used in Exp. I, except for a change in the critical portions of the instructions. For each pair of displays, *O* was asked to decide which display showed the greater coherence, with coherence defined as "the degree to which the parts of the display seem to maintain the same relative positions as the display moves." No mention was made of the possibility of seeing the displays either as two-dimensional or three-dimensional.

Results

Each treatment combination was represented in 19 paired comparisons. For each of the 20 *O*s, the number of times a stimulus representing each treatment combination was chosen[3] was calculated. The method employed to analyze these frequencies is a direct extension to three factors (numerosity, perspective, and *O*s) of the median test described by Mood (1950) for "two factor experiments, one observation per cell" (pp. 399–402).

In Exp. I significant effects ($P < .05$) were found for numerosity ($\chi^2 = 125.96$, $df = 4$), perspective ($\chi^2 = 21.90$, $df = 3$), and the interaction of the two treatment dimensions ($\chi^2 = 108.30$, $df = 19$). The proportion of trials on which stimuli representing each of the treatment combinations were chosen by all *O*s is given in Table 1. A greater effect of perspective is indicated for the stimuli having smaller numbers of points. The relationship between the propor-

[3] "Chosen" is used to mean "chosen as giving a stronger impression of occurring in space than the stimulus with which it was paired" when reference is made to Exp. I, and "chosen as appearing more coherent than the stimulus with which it was paired" when reference is made to Exp. II.

TABLE 1

PROPORTION OF TRIALS ON WHICH STIMULI WERE CHOSEN AS GIVING A STRONGER DEPTH IMPRESSION: EXP. I

Perspective	Number of Points					All Numbers
	2	3	4	5	6	
1.00	.232	.329	.445	.524	.561	.418
1.13	.245	.508	.547	.537	.632	.494
1.67	.358	.376	.621	.634	.589	.516
3.00	.447	.524	.589	.608	.695	.573
All levels	.320	.434	.551	.576	.619	

TABLE 2

PROPORTION OF TRIALS ON WHICH STIMULI WERE CHOSEN AS APPEARING MORE COHERENT: EXP. II

Perspective	Number of Points					All Numbers
	2	3	4	5	6	
1.00	.708	.598	.517	.665	.541	.606
1.13	.685	.594	.517	.527	.603	.585
1.67	.541	.522	.483	.503	.483	.506
3.00	.527	.273	.287	.230	.201	.303
All levels	.614	.496	.451	.481	.457	

tion of times stimuli representing each level of numerosity were chosen and the logarithm of the number of points for that level, is illustrated in Fig. 2. Figure 3 illustrates the corresponding relationship for levels of perspective.

The frequencies of trials on which stimuli representing each treatment combination were chosen by each of the 11 *O*s in Exp. II were also subjected to a median analysis. Significant effects ($P < .05$) were found for numerosity ($\chi^2 = 25.59$, $df = 4$), perspective ($\chi^2 = 42.27$, $df = 3$), and their interaction ($\chi^2 = 63.22$, $df = 19$). The proportion of trials on which stimuli representing each of the treatment combinations were chosen by all *O*s is given in Table 2. The relationships between numerosity and proportion of choice and between perspective and proportion of choice are illustrated in Fig. 2, and Fig. 3, respectively.

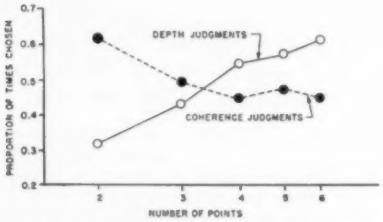

FIG. 2. Effect of numerosity on depth and coherence judgments.

Figures 2 and 3 also illustrate the opposite effects of the treatment dimensions on the two dependent variables, depth judgments and coherence judgments. As well as difference in direction of effect, the treatment dimensions showed differences in relative importance in their effects on the two types of judgments. The 20 proportions in Table 1 are ordered primarily on the basis of numerosity and secondly on the basis of perspective, while the proportions in Table 2 are primarily ordered on perspective and secondly on numerosity.

DISCUSSION

The proportion of trials on which a stimulus was judged to give a stronger impression of occurring in space than the stimulus with which it was paired was found to be an approximately linear function of the logarithm of the number of points in the stimulus. This supports

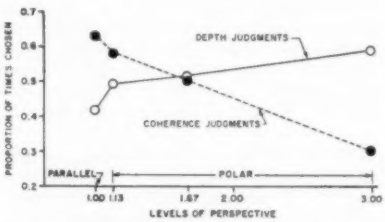

FIG. 3. Effect of perspective on depth and coherence judgments.

the hypothesis of a direct relationship between number of points and judged strength of the depth impression. Although the logarithmic relationship was not specifically predicted, it is the same as that obtained by Green (1959a) using larger numbers of spots and ratings of subjective coherence.

Perspective significantly affected the choice of stimuli giving stronger depth impressions, although it appeared to be secondary to number of points in its effect on the ordering of the stimuli. This would indicate that the interrelated factors of motion parallax and linear perspective do affect the depth impression created by these displays.

The question arises as to whether number of points may be important only because increasing numbers of points provide additional opportunities to view the effects of perspective. This is contraindicated by the findings that larger differences between levels of perspective occurred for smaller numbers of points and that the relationship between number of points and proportion of choice holds even for parallel perspective. It would appear instead that perspective becomes more important as the cues to change in orientation of the pattern are reduced, i.e., with smaller numbers of points.

Some insight into the role of subjective rigidity or coherence of the patterns in influencing perceived depth may be gained from the results of the second experiment, which showed different effects of the independent variables when coherence judgments were used as the dependent variable. Displays of two spots were judged most coherent. This would be expected, as two spots moving on a plane could always represent the projections of two points at a constant distance in three dimensions. The relationship between numerosity and coherence judgments for larger numbers of spots was not very pronounced.

A much greater effect on coherence judgments was produced by perspective. Stimuli produced with closer viewing points were judged less coherent than those with more distant viewing points. There was thus an inverse relationship between depth judgments and coherence judgments across levels of perspective. This would seem to result from a combination of two factors. First, as all stimuli were not perceived as three-dimensional at all times, the changes in distances between the two-dimensional projections of the points may have influenced average coherence judgments, and more variability in the two-dimensional projections did occur with greater perspective. The other factor is the distortion resulting from the magnification and demagnification of the projections used to maintain a uniform "average" size. This would especially affect the stimuli with maximum perspective, which differed furthest from appropriate perspective, and the difference between depth and coherence judgments is greatest in this case.

The use of perspective has been an area of disagreement between J. J. Gibson and H. Wallach, the two major experimenters in this area. Gibson (1957) uses what he properly regards as appropriate perspective in his displays. A translucent screen is placed *equidistant* between *O* and a point light source and shadows of objects rotating between the light source and screen are observed. This is equivalent to displaying projections computed with the projection point at the position at which *O* will place his eye when he observes the stimuli. Wallach (1953), on the other hand, uses a distant projection point which is further from the screen than *O*'s eye. The resulting situation is similar to the use of parallel perspective in the present research.

Gibson's situation is more like the real-life perceptual situation and present results would indicate that it should in general lead to a stronger impression of depth than does Wallach's method. The "kinetic depth effect," as formulated by Wallach, does however require parallel projections of the three-dimensional displays for its isolated study.

The choice of perspective should then be made according to the purpose of the experimenter in these respects. The

present results indicate that the importance of variations in perspective will depend upon the type of judgments used. When the effects of variation in perspective are of interest, at least for displays involving small numbers of elements, depth judgments and rigidity judgments cannot be assumed equivalent.

Summary

Motion picture sequences of spots representing projections of points rotating in three dimensions were produced using the CRT output of a digital computer. The sequences varied in number of points and in perspective. Perspective is defined as the ratio of the distance from the projection point to the most distant plane of the display to the distance from the projection point to the closest plane of the display. A paired comparison method was used to elicit judgments of the relative strength of the depth impressions created by the sequences and of the relative coherence of the patterns while in motion.

Judged strength of the depth impression increased with increasing numbers of spots and, to a lesser degree, with increasing perspective. Subjective coherence decreased with increasing perspective. The role of perceived coherence of moving patterns in depth perception and the effects of variations in perspective on depth judgments are discussed.

References

Gibson, J. J. *The perception of the visual world.* Boston: Houghton Mifflin, 1950.

Gibson, J. J., & Gibson, E. J. Continuous perspective transformations and the perception of rigid motion. *J. exp. Psychol.*, 1957, **54**, 129–138.

Green, B. F. Kinetic depth effect. (Psychology Group 58) Quarterly Progress Report, Massachusetts Institute of Technology, Lincoln Laboratory, 1959. (a)

Green, B. F. Mathematical notes on 3-D rotations, 2-D perspective transformations, and dot configurations. Group Report No. 58-5, Massachusetts Institute of Technology, Lincoln Laboratory, 1959. (b)

Metzger, W. Tiefenerscheinungen in optischen Bewegungsfeldern. *Psychol. Forsch.*, 1934, **20**, 195–260.

Miles, W. R. Movement interpretations of the silhouette of a revolving fan. *Amer. J. Psychol.*, 1931, **43**, 392–405.

Mood, A. M. *Introduction to the theory of statistics.* New York: McGraw-Hill, 1950.

Wallach, J., & O'Connell, D. N. The kinetic depth effect. *J. exp. Psychol.*, 1953, **45**, 205–217.

(Received October 2, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 421

SUPPLEMENTARY REPORT: PROACTIVE INHIBITION AS A FUNCTION OF THE METHOD OF REPRODUCTION [1]

JOHN L. WIPF AND WILSE B. WEBB

University of Florida

Greenberg and Underwood (1950) found that paired-associate recall, after 5 hr., was inversely related to the number of lists previously learned. Thus, proactive inhibition occurred. Their learning materials were two-syllable adjectives arranged in four paired-associate lists of 10 pairs each. The anticipation method was used to test recall.

Method.—The present study is a replication of the Greenberg and Underwood study in that it employed a 5-hr. retention interval, a 2-sec. presentation rate on a memory drum, the same learning-relearning criteria (8/10 correct and 10/10 correct, respectively), and four lists over 4 days. However, in the present study, the method of presentation was modified to permit the use of free recall or reproduction as a retention measure. The 10 Greenberg-Underwood response words in a list were shown one at a time. After each trial (one presentation of the 10 words), *Ss* were required to reproduce in writing the words they could remember. A 3-min. limit was set on this recall. Successive trials, using four different serial arrangements, were given until learning criterion was met.

The method of reproduction was similarly employed in testing recall, i.e., after 5 hr. *Ss* reproduced, in writing, as many words as they could recall. Recall was immediately followed by relearning. A different list was used on each of 4 successive days. Twenty-four different orders of the four lists were used. Twenty-four *Ss* (11 men and 13 women introductory psychology students) each learned the lists in a different order.

Results.—Separate analyses of three dependent measures, number of trials to learn, number of words correctly recalled, and number of trials to relearn, were performed. The means and *SD*s of these measures are given in Table 1. An analysis of variance showed the differences among the mean number of words correctly recalled over days to be significant ($P < .001$). Thus, significant PI at recall occurred from day to day, after the first day. Recall differences among *Ss* were significant. Recall differences among lists were not significant.

TABLE 1

MEANS AND *SD*s OF TRIALS TO LEARN, WORDS RECALLED, AND TRIALS TO RELEARN

Day	Trials to Learn		Words Recalled		Trials to Relearn	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
1	2.46	.76	6.17	2.15	3.00	3.27
2	2.91	1.55	4.95	2.42	2.79	2.25
3	2.58	.81	4.62	2.64	2.17	1.11
4	2.21	.91	3.92	2.44	2.58	1.71

The analysis of the learning measures showed differences due to days to be significant ($P < .05$). These differences appear to result from facilitation effect (see Table 1 means, Days 2 to 4). List differences were significant ($P < .01$). Thus, word lists differed in difficulty in terms of trials-to-learn. Subject differences were again significant.

List difficulty in terms of trials-to-learn should have had no differential effect because all possible sequences of the four lists were used; hence list differences were counterbalanced throughout the study.

The analysis of the relearning measures showed Day and List differences not to be significant. Subject differences were significant.

Allowing ample time for response and written reproduction, the use of unpaired presentations increased the number of words recalled and decreased the number of trials required to learn and relearn (compare Table 1 with the Greenberg-Underwood results). However, a PI effect was demonstrated. A minimal conclusion which may be drawn from these data is that the PI effect demonstrated by Greenberg and Underwood is not a function of the particular learning-recall procedure utilized.

[1] Based on a thesis submitted to the Graduate School, University of Florida, in partial fulfillment of the requirements for the Master of Arts degree.

REFERENCE

GREENBERG, R., & UNDERWOOD, B. J. Retention as a function of stage of practice. *J. exp. Psychol.*, 1950, **40**, 452–457.

(Received July 13, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 422–423

SUPPLEMENTARY REPORT: EFFECT UPON SENSORY PRECONDITIONING OF BACKWARD, FORWARD, AND TRACE PRECONDITIONING TRAINING[1]

JAMES D. WYNNE AND W. J. BROGDEN

University of Wisconsin

Hoffeld, Thompson, and Brogden (1958) report variation in magnitude of sensory preconditioning (SP) in cats as a function of stimuli time relations during preconditioning. The 200 preconditioning training trials involved a 2-sec. light paired with a 1000-cycle tone either of 2, 2.5, 3, 4, or 6 sec. duration. Since tone and light terminated together, the independent variable of tone precedence is represented by values of 0, .5, 1, 2, and 4 sec. Significant linear and quadratic trend was found for magnitude of SP with the maximum at 4 sec. precedence. There is, however, a confounding of precedence and duration, and the longer durations of tone represent conditions for delayed conditioning. If a fixed duration of tone is used, there is no confounding of duration with precedence, but the large values of precedence then represent the conditions of trace conditioning.

Method.—The present experiment is a replication of the experiment reported by Hoffeld et al. (1958) except for the following differences: (*a*) duration of the tone was always held constant at 2 sec.; (*b*) the range of precedence of tone was increased and included negative values (backward conditioning) because Silver and Meyer (1954) report equal magnitudes of SP for simultaneous and a backward preconditioning time relation; (*c*) the number of preconditioning training trials was reduced from 200 to 4 because Hoffeld, Kendall, Thompson, and Brogden (1960) found maximum SP for the latter value; (*d*) N for each group was 8; (*e*) one session of adaptation to the rotator was introduced; and (*f*) following the test for SP all *S*s were given tone-shock avoidance training in the attempt to obtain an additional measure of SP. Since this procedure produced no significant results, no report of them is made.

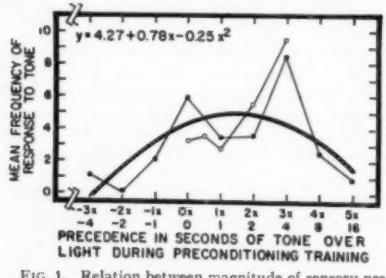
y = 4.27 + 0.78x − 0.25 x²
MEAN FREQUENCY OF RESPONSE TO TONE
0
2
4
6
8
10
−3s −2s −1s 0s 1s 2s 3s 4s 5s
−4 −2 −1 0 1 2 4 8 16
PRECEDENCE IN SECONDS OF TONE OVER LIGHT DURING PRECONDITIONING TRAINING

FIG. 1. Relation between magnitude of sensory preconditioning and precedence of tone over light during preconditioning. (The data represented by the open circles is taken from Hoffeld et al., 1958. The filled circles represent the data of the present experiment. The smooth curve is the best fitting quadratic function for the data of the present experiment, the equation for which is given in the upper left corner of the figure.)

Results.—Magnitude of SP is measured by frequency of response in the cross-modal generalization tests of tone alone, given following avoidance conditioning to the light CS. An orthogonal polynomial analysis of variance (Grant, 1956) of the data for the experimental groups shows significant differences between groups and a significant quadratic trend ($P = .05$). The experimental group means are plotted in Fig. 1 as is also the curve of the best-fitting quadratic equation. Range tests (Duncan, 1951) indicate that the mean for 4 sec. precedence is significantly different from all other means, none of which differ significantly from each other. The means for precedence conditions −4, −2, and 16 sec. are either approximately equal to or less than the mean for the control group (.9), so it is improbable that any SP occurs for these conditions.

The control group of the present study shows greater cross-modal generalization than similar groups have shown in earlier studies (Brogden, 1939; Hoffeld et al., 1958). If the control *S*s of Hoffeld et al. (1958) are added to the present control group, N is 12 and the mean is .58. The upper fiducial limit for cross-modal generalization provided by the t test is .83 at the 5% level. This limit clearly excludes evidence of SP at precedence levels of −4 and −2 sec. (backward conditioning) and long trace conditioning at 16 sec. precedence. If the combined control group data are used in one-tailed t tests with that of the remaining experimental groups, the means for precedence conditions 0, 1, 2, and 4 sec. exceed the control group mean at the 5% level, and for precedence conditions −1 and 8 sec.

[1] This research was supported in part by grants from the National Science Foundation and the Research Committee of the Graduate School from funds provided by the Wisconsin Alumni Research Foundation.

exceed the control mean at the 20% and 10% levels, respectively.

Discussion.—The present results are in good agreement with those of Hoffeld et al. (1958) which are plotted in Fig. 1. In comparing the two studies it is evident that the confounding of duration of tone with its precedence ove light during preconditioning training has little if any effect upon the magnitude of SP. Saying it another way, delay and trace conditioning procedures during preconditioning make little if any difference as long as the time relations between the onset of the CS (tone) and onset of the UCS (light) are the same.

A difference in magnitude of SP, favoring the present experiment, was expected on the basis of parametric differences in the number of preconditioning trials, since Hoffeld et al. (1960) found this variable to have a marked effect upon the magnitude of SP. They obtained a mean frequency of response to trials of tone alone of 24.17 for 4 preconditioning trials (value used in the present study) and a mean of 10.50 responses for 200 trials (value used by Hoffeld et al., 1958). The time relations during preconditioning of the Hoffeld et al. (1960) experiment were those found optimal by Hoffeld et al. (1958) and involved a 6-sec. tone and 2-sec. light with 4 sec. precedence. In view of the minor differences in amount of SP for comparable time relations between the present study and that of Hoffeld et al. (1958), the possibility of interactions between trace and delayed time relations with number of preconditioning trials must be considered. The possibility also exists that an unknown variable or parameter present in the Hoffeld et al. (1960) study accounts for the high level of SP.

Although a significant quadratic trend in magnitude of SP as a function of time relations was found in the present experiment, the best-fitting quadratic equation does not give a satisfactory description of the relationship in the data. Other analyses establish a maximum effect at tone precedence of 4 sec., with no effect for the backward conditions of −4 and −2 sec. precedence or for the long trace condition of 16 sec. precedence. Whether SP occurs at the backward condition of −1 sec. precedence and the trace condition of 8 sec. precedence is dubious. These two conditions represent the extremes between which fall the time relations during preconditioning that are effective in producing SP. A quadratic function with a maximum around 4 sec. precedence and zero at the backward condition of approximately −1-sec. precedence and the forward condition of approximately 8 sec. precedence may be reasonable, but is not clearly supported by the data of the experiment.

References

Brogden, W. J. Sensory preconditioning. *J. exp. Psychol.*, 1939, **25**, 323–332.

Duncan, R. B. A significance test for differences between ranked treatments in an analysis of variance. *Va. J. Sci.*, 1951, **2**, 171–189.

Grant, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.

Hoffeld, D. R., Kendall, S. B., Thompson, R. F., & Brogden, W. J. Effect of amount of preconditioning training upon the magnitude of sensory preconditioning. *J. exp. Psychol.*, 1960, **59**, 198–204.

Hoffeld, D. R., Thompson, R. F., & Brogden, W. J. Effect of stimuli time relations during preconditioning training upon the magnitude of sensory preconditioning. *J. exp. Psychol.*, 1958, **56**, 437–442.

Silver, C., & Meyer, D. R. Temporal factors in sensory preconditioning. *J. comp. physiol. Psychol.*, 1954, **47**, 57–59.

(Received July 13, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 4, 423–424

SUPPLEMENTARY REPORT: SEMANTIC GENERALIZATION IN PROBABILITY LEARNING

J. P. DAS

Utkal University, Cuttack, India

This experiment is planned to find out whether the introduction of semantic stimuli changes the course of generalization in probability learning. The present study is similar in design and procedure to one by Popper and Atkinson (1958). Both are based on predictions from a discrimination model and its application given in Estes and Burke (1955). Every trial begins with the presentation of a T_1 or a T_2 stimulus, the probability of each being .50. Following a T_1 stimulus, the letter X occurs with a probability of .90 and Y with a probability of .10. But following a T_2 stimulus, X occurs with a probability of .70 for one group and .30 for another group; the corresponding probability of Y being .30 and .70, respectively. These probability schedules are identical with those of Groups II and IV of a related experiment by Atkinson, Bogartz, and Turner (1959).

TABLE 1

MEANS AND *SD*s OF X PREDICTIONS OVER 20 TRIALS FOLLOWING T_1 AND T_2

Groups		Trial Blocks							
		1		2		3		4	
		Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
I	R90	13.53	2.53	17.30	1.80	18.40	2.18	18.73	1.61
	D70	11.13	3.32	12.93	2.82	15.00	3.44	15.27	2.93
II	R90	13.86	2.29	15.38	2.42	17.29	2.33	17.86	2.17
	D30	8.62	1.89	7.29	3.04	4.62	2.34	3.33	2.71
III	R90	14.54	4.52	17.38	2.27	17.85	1.70	17.92	2.23
	N70	12.15	3.28	14.85	2.82	17.15	1.56	17.46	1.15
IV	R90	13.38	2.27	14.38	3.20	15.85	3.30	16.46	3.22
	N30	9.69	2.64	8.69	3.87	6.46	3.63	5.38	3.58

Note.—t for R in I and III = 1.04, df = 26; for R in II and IV = 1.33, df = 32; for D and N in I and III = 2.58, df = 26, $P < .02$; for D and N in II and IV = 1.71; df = 32.

However, there is a major difference between the Popper and Atkinson experiment and the present one. Synonyms and antonyms have been used here in place of nonsense syllables as T_1 and T_2 stimuli.

Method.—The T_1 stimulus was an Indian word, RAJANI, meaning night. The two T_2 stimulus words, NIGHT and DAY, were synonymous and antonymous to RAJANI. When stimulus words were visually presented, they were written in the common script of *Ss*, Oriya. There were four experimental groups: I, R90 D70 (N=15); II, R90-D30 (N=21); III, R90 N70 (N=13); and IV, R90-N30 (N = 13). R90-D70 stands for the group in which T_1 was RAJANI and probability of X after T_1 was .90, whereas T_2 was DAY and the probability of X after T_2 was .70. The other terms are similarly defined.

Each group of *Ss* received 160 trials in all, divided into four 40-trial blocks. Within each block, the 20 T_1 and 20 T_2 trials were randomly distributed, and in the T_1 and T_2 blocks, the given proportions of X and Y were maintained. For example in Group I, out of the 20 T_1 cards, the combination RAJANI-X was found in 18 cards whereas the remaining 2 cards had RAJANI-Y. Similarly the 20 T_2 cards had 14 cards of DAY-X and 6 of DAY-Y.

The *Ss*, college students, had four strips of paper for the four blocks of trials. On each strip they found two columns, marked with T_1 and T_2 stimuli at the top. Their task was to guess whether X or Y would come and record it under the appropriate columns immediately after *E* read out the T_1 or T_2 stimulus. Then *E* exposed the card, and at the same time read out whatever was written following the T_1 or T_2 stimulus. Each trial took about 10 sec.: 3 sec. to read out the stimulus word, 4 sec. allowed to *Ss* for putting down their guesses, and 3 sec. to expose the card to *Ss* and begin the next trial. The *Ss* covered their previous choices by folding the recorded portion of the paper after each trial.

Results.—A summary of the findings is given in Table 1. The *Ss*' predictions were expected to vary on two accounts—semantic relation between T_1 and T_2, and similarity between the probabilities of X after T_1 and T_2. Groups I and III, and Groups II and IV therefore, may be meaningfully compared. The final proportion of R choices in Group I should be higher than the same in Group III because there will be a greater generalization between the synonyms R and N than between the antonyms R and D. Similarly, the mean prediction of R in the final trial block for Group IV will be less than that for Group II, whereas the reverse would be true for N and D. The series of t's computed shows that although the mean differences are in the expected direction, only the difference between N70 and D70 reached at least the .05 level of significance. This result may be interpreted as offering limited support to the hypothesis that semantic stimuli influence probability generalization.

REFERENCES

ATKINSON, R. C., BOGARTZ, H. W., & TURNER, R. N. Discrimination learning with probabilistic reinforcement schedules. *J. exp. Psychol.*, 1959, **57**, 349-350.
ESTES, W. K., & BURKE, C. J. Application of a statistical model to simple discrimination learning in human subjects. *J. exp. Psychol.*, 1955, **50**, 81-88.
POPPER, J., & ATKINSON, R. C. Discrimination learning in a verbal conditioning situation. *J. exp. Psychol.*, 1958, **56**, 21-25.

(Received August 14, 1961)

A Word to the Student

You can save money by beginning your personal file now. Current subscriptions for a given year, beginning with the January issue, are available to students at a special rate of $6.50. This is a saving of $5.50 per volume. By acting early, you can obtain both Volume 5 (1962) and Volume 6 (1963) at this rate. If bought later, as back volumes, they will cost $12.00 each.

JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR

Please address all orders and business inquiries to:

Kay Dinsmoor, JEAB
Department of Psychology
Lindley Hall
Indiana University
Bloomington, Indiana

Checks should be made payable to the journal.

Two Fascinating New Books for Psychologists

★ PERCEPTION AND MOTION

By Karl U. Smith, Ph.D., University of Wisconsin and William M. Smith, Ph.D., Dartmouth College. About 432 pages, about 193 illustrations—Just Ready!

New answers to question of how patterned motion is regulated

Here is a most unusual survey of the problems of behavior organization in man—considered in terms of sensory feedback mechanisms, both visual and auditory. The authors have developed a new theory toward explaining certain aspects of behavior, perception and motor learning. Their neurogeometric theory revolves around the concept that the brain is a differential detector responding to differences in stimulation and sensory feedback—that the differential action within the nervous system is based on varying input to the neuron from several sources of stimulation rather than being based on the decision-making powers of the synapse. Other elements of the theory hinge on perception and motion being considered as a single mechanism rather than as two separate forms of adaptation; and on motion itself being divided into postural, transport and manipulative components—each regulated by particular stimulation patterns.

Research techniques used to develop and support the authors' contentions largely involve electronic methods of motion analysis and use of closed circuit television and videotape recordings.

Findings and conclusions cast doubt upon many widely accepted concepts about the reflex, the synapse, and the nature of the learning process (particularly on the validity of the reinforcement theory). To supply background information for viewing their work in proper perspective, the authors briefly review historical development of the science of motion.

Contents embrace the following discussions: Concepts of Perceptual-Motor Organization—The Science of Motion—Motion, Work and Machines—Early Experiments of the Spatial Problems of Behavior—Recent Studies of Displaced Vision—Critical Summary of Displaced Vision Studies—Experimental Foundations of Neurogeometric Theory—Use of Television in Behavior Research—Televised Inversion and Reversal of Visual Feedback—Televised Angular Displacement of Visual Feedback—Size-Distorted Visual Feedback—Displaced Vision and Theory of Tool Using—Delayed Sensory Feedback—Infant Control of the Behavioral Environment—Effects of Displaced and Delayed Feedback—Neurogeometry of Motion.

★ DELAYED SENSORY FEEDBACK AND BEHAVIOR

By Karl L. Smith, Ph.D. About 128 pages, illustrated—Just Ready!

A new type of temporal analysis of behavior

This book is an elaboration of a portion of the book described above. It investigates the strange effects produced by an artificially induced time lag between performance of motion and perception of that motion. This investigation bears directly upon present day theories of learning and behavior; and gives new insight into human operation of complex machine systems. In this study, through use of a closed television circuit, the performer's actions are fed back to him pictorially—but have been displaced spatially by manipulating the electronic circuits, the camera lenses, or the camera itself; or the image has been delayed by means of videotape recordings.

Contents are divided into six broad sections: Significance of Delayed Sensory in Behavior Research—Delayed Visual Feedback in Tracking Performance—Delayed Auditory Feedback—Blind Performance with Delayed Pictorial Feedback—Concurrent Delayed Visual Feedback Implications of Experiment.

Under these, are discussed such unusual topics as delayed graphic feedback of handwriting, artificial stutter, delayed guidance feedback from remote cosmic systems: theoretical considerations.

Teachers and students of psychology and education; industrial psychologists; certain researchers in television techniques; time and motion engineers will all find valuable and thought-provoking information in these two new books.

W. B. SAUNDERS COMPANY, West Washington Square, Philadelphia 5, Pa. JLEP

Please send me the following books and bill me:

☐ Smith and Smith—**Perception and Motion,** about $9.00.

☐ Smith—**Delayed Sensory Feedback and Behavior,** about $6.00.

Name.....**Address**.....